



National Institute of Child Health and Human Development

NICHD

ANNUAL REPORT

FY95

October 1, 1994 to September 30, 1995

PART 4

RJ
85
N277
1995
pt. 4

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 00066-25 LMG

PERIOD COVERED

October 1, 1994 through September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Control Mechanisms in Temperate Bacteriophage λ

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	Robert A. Weisberg	Section Head	LMG:NICHD
Others:	Banik-Maiti, Sarbani	Visiting Fellow	LMG:NICHD
	Dorgai, Laszlo	Visiting Associate	LMG:NICHD
	King, Rodney	NRC Fellow	LMG:NICHD
	Mzhavia, Nino	Visiting Fellow	LMG:NICHD
	Sloan, Sieghild	Microbiologist	LMG:NICHD
	Yagil, Ezra	Prof. Serv. Contract	LMG:NICHD

COOPERATING UNITS (if any)

Department of Biochemistry, Tel Aviv University, Israel (Ezra Yagil); Laboratory of Molecular Biology, NCI (Ding Jun Jin)

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Microbial Genetics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

6.2

PROFESSIONAL:

5.2

OTHER:

1.0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Antitermination increases the transcription of genes that are located downstream of terminators. Phage HK022 antiterminates transcription of its early genes by converting host RNA polymerase to a terminator-resistant form. This requires cis-acting phage sequences called nut sites. The isolation and characterization of nut region mutants argue that the sites are about 70 bases long and act as RNA transcripts consisting of two adjacent stem-loops. These transcripts presumably interact with a recognition site on RNA polymerase. The properties of polymerase mutants that are unable to antiterminate HK022 early transcription but are normal in all other respects tested suggest that the recognition site is located in a zinc-binding domain of the β' subunit. The modified polymerase is capable of reading through at least three successive strong transcription terminators with much increased efficiency. The efficiency is greater than 90% per terminator in vivo and about 50% in vitro, consistent with the hypothesis that the requirements for persistent antitermination in the HK022 system are extraordinarily simple. The integrase proteins of phages λ and HK022 are closely related site-specific recombinases that recognize different nucleotide sequences in the core regions of their substrates, the attachment sites of the two phages. The two proteins differ by 92 amino acid substitutions. We have found that no more than 5 of these substitutions are critical for the different DNA sequence specificities of these enzymes. Two of these substitutions act principally by reducing specificity, while the other three act principally by specifically decreasing activity on one of the substrates.

Project Description:

Objective: To understand the mechanisms of genetic recombination, especially as exemplified by the insertion of the bacteriophage λ genome into the *Escherichia coli* chromosome, and to understand how viruses control the expression of their genes.

Major Findings.

1. Antitermination of early transcription in phage HK022.

We have previously shown that HK022 antiterminates early transcription, but that it does so in a way that differs from and appears far simpler than previously characterized antitermination systems. The only specific requirements for antitermination appear to be cis-acting phage sites called *nutL* and *nutR*, located in the HK022 P_L and P_R operons, respectively, and a protein region (the "zinc-binding") located near the amino-terminal end of the β' subunit of *E. coli* RNA polymerase. The numerous protein factors needed in other antitermination systems are not required by HK022. In order to define the *nut* sites and to ask if they interact with the zinc-binding region of β' , we have developed a genetic screen to isolate and to characterize mutants that alter antitermination. We fused the HK022 P_L promoter and early transcribed sequences to a promoterless *lac* operon. Three strong tandem transcription terminators can be inserted between the promoter and the *lac* operon (Figure 1). These terminators have no effect on β -galactosidase production if they are preceded by a sufficiently long segment of the HK022 early transcribed sequences and if the zinc binding region of RNA polymerase has the wild type sequence, but they reduce the production of β -galactosidase to about 0.2% if one of these conditions is not met.



Figure 1. The promoter-reporter gene fusion used to isolate and characterize antitermination mutants.

Using this system, we found that antitermination is prevented or strongly reduced by linker scanning and deletion mutations that alter the first 75 NT (approximately) of the wild type P_L transcript. This region, called *nutL*, is similar in sequence to a region that begins at position 277 of the HK022 P_R transcript (*nutR*). The predicted secondary structures of the RNA transcripts of both segments are also similar, consisting of two adjacent stem-loops (Figure 2), and substitution of *nutR* for *nutL* in the P_L -*lacZ* fusion does not change the efficiency of antitermination. Therefore, *nutR* is functionally equivalent to *nutL*. To see if the predicted secondary structure is functionally important, we isolated *nutL* mutations that disrupted the base pairs indicated by "G" in Figure 2. A single change in stem 1 or a double change in stem 2 reduced antitermination, and both changes together completely abolished it (i.e., terminator readthrough was similar to that seen in a *nutL* deletion). However, antitermination was completely restored if compensatory changes that restored the predicted base-pairing (but not the original sequence) were introduced. These findings argue that the proposed secondary structure is functionally important and that the *nut* sites act, at least in part, as RNA transcripts.

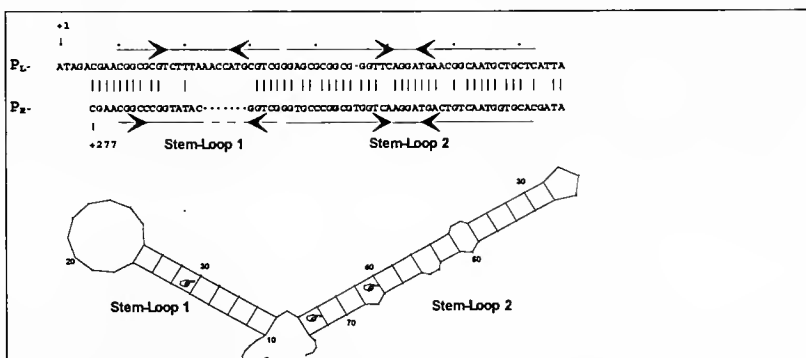


Figure 2. Top: The *nutL* (first line) and *nutR* (second line) sites of HK022 are aligned so as to maximize sequence and secondary structure similarities. The numbers +1 and +277 indicate position relative to the starts of the P_L and P_R transcripts, respectively. Bottom: Base-pairings in the predicted *nutL* structure are shown. "G" marks base-pairs that vary between *nutL* and *nutR* in such a way as to preserve pairing.

We have made additional changes to assess the importance of non-base-paired regions of the proposed structure. A change of the base between the two stem-loops (G35A) destroyed *nutL* function as did base substitutions or deletions

that altered the internal loop between positions 44 and 47. By contrast, a base substitution in the loop of stem-loop 2 had no effect on function.

2. Antitermination with purified proteins (in collaboration with D. Jin).

In vivo the HK022 *nut* sites promote efficient readthrough of several sequential termination signals. The most promoter-distal terminator can be 5 kbp or more from the anti-termination site. This argues that the *nut* sites participate in conversion of transcribing RNA polymerase to an antiterminating form and that this form persists for considerable time and distance. Does such persistent antitermination require proteins other than polymerase itself? To answer this question, we transcribed wild type and mutant *nut* templates with purified wild type or mutant RNA polymerases. These templates are similar to the triple terminator construct depicted in Figure 1 except that the DNA was cleaved downstream of the third terminator. We found that purified wild type polymerase read through three successive terminators with an overall efficiency of 15 to 26% (approximately 50% per terminator). Figure 3 shows a phosphorimager scan of a gel in which the products of such a reaction, labeled with α - P^{32} -CTP, were fractionated. Mutations of the zinc-binding region of the β' subunit (*rpoCY75N*) or of the *nutL* site, either of which blocks antitermination *in vivo*, reduced the efficiency of readthrough of the three terminators to nearly undetectable levels (10% or less per terminator). We have obtained similar or identical results with polymerases purified by two different independent methods: adsorption of histidine-tagged enzyme onto nickel-agarose and reconstitution of polymerase from denatured subunits. We conclude that this *in vitro* system for measurement of terminator readthrough reflects antitermination efficiencies measured *in vivo*, and that no protein other than RNA polymerase is required for HK022-specific antitermination. However, we note that readthrough *in vitro* is less efficient than it is *in vivo*, and it therefore remains possible that additional proteins stimulate or control the level of antitermination.

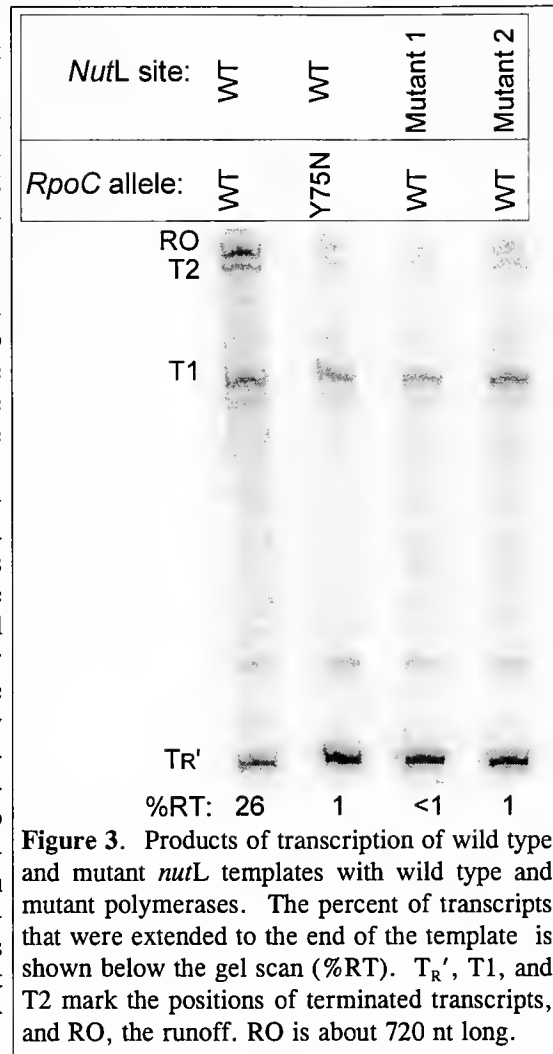


Figure 3. Products of transcription of wild type and mutant *nutL* templates with wild type and mutant polymerases. The percent of transcripts that were extended to the end of the template is shown below the gel scan (%RT). T_R' , T1, and T2 mark the positions of terminated transcripts, and RO, the runoff. RO is about 720 nt long.

3. Recognition of attachment sites by integrase proteins (in collaboration with E. Yagil).

The Integrases of bacteriophages λ and HK022 promote recombination between DNA molecules that carry attachment sites. The two integrases are about 70% identical in sequence and catalyze nearly identical reactions, but recognize different sets of sites. To identify the amino acids that determine this difference in specificity, we selected mutants of λ integrase with increased ability to recombine HK022 sites. This selection yielded eleven different amino acid substitutions at eight different positions. Three of the positions belong to a larger set that were identified as important for the λ /HK022 specificity difference by analysis of chimeric integrases. Substitution of the HK022 for the corresponding λ residue at each of these three positions increased recombination of HK022 sites, and one double substitution, N99D-E319R, increased recombination to nearly wild type HK022 levels. Mutations at the other five positions changed residues that are identical in the wild type proteins or are at positions identified by chimera analysis as unimportant for the λ /HK022 specificity difference. All of the mutants isolated by selection for increased recombination of HK022 sites retained considerable ability to recombine λ sites. However, we found that substitution of HK022 for λ residues at three additional positions, S282P, G283K, and R287K, specifically reduced recombination

of λ sites. These three substitutions when combined with N99D and E319R were sufficient to change the specificity of λ to that of HK022 integrase. The first three substitutions act principally to prevent recombination of λ sites, and the second two to remove a barrier to recombination of HK022 sites. We suggest that many natural alterations in the specificity of protein-DNA interactions occur by multi-step changes that first relax and then restrict specificity.

4. Identification of nucleotides that are differentially recognized by HK022 and λ integrases.

The HK022 and λ integrases recognize different sequences within the core binding sites of their respective attachment sites. Each attachment site has two core binding sites arranged as pair of inverted repeats separated by a spacer sequence. DNA strand exchange between attachment sites occurs within the spacer. The phage and bacterial attachment sites of HK022 and λ have 4 core binding sites (called B, B', C, and C'), and each differs from the others in nucleotide sequence. We have constructed recombination reporter substrates to help us to determine which nucleotides of these sites determine the HK022/ λ specificity difference (Figure 4). Int-promoted excision of the segment between two attachment sites (denoted B'-C and C'-B in Figure 4) allows the production of active β -galactosidase. Our results to date suggest that the λ and HK022 B' sites contain important determinants of specificity: substrates whose core sites consist of 4 λ B' sequences or 4 HK022 B' sequences have the corresponding recombination specificity. By contrast, sites with 4 λ C' sequences can be efficiently recombined by both integrases. Since the C' sites of the two phages have almost the same sequence, this result is not surprising and confirms a previous hypothesis that the two integrases have overlapping rather than completely different specificities.

5. The action of the Nun transcription-termination factor *in vivo*.

The HK022-encoded Nun protein prevents transcript elongation in regions lying promoter-distal to the *nut* sites of phage λ . Although this can be formally described as transcription termination, in fact it is difficult to distinguish between transcription termination and transcription arrest *in vivo*. Both processes produce new 3' RNA ends, and both prevent the transcription of downstream genes. However, an arrested transcription complex should physically block the completion of subsequent rounds of transcription, whereas termination should not block subsequent rounds of transcription because the complex will have dissociated from the template. Dr. M. E. Gottesman and his collaborators (personal communication) have found that Nun arrests rather than terminates transcription *in vitro*. To detect arrested transcription complexes *in vivo*, we have measured the effect of Nun on the level of *nut*-proximal transcript in the P_L operon of phage λ . We used two methods to avoid the problems that could be caused by an effect of Nun on transcript stability. One set of experiments was carried out in a multiply mutant strain that is unable to produce several of the nucleases that are responsible for message degradation in *E. Coli*. These mutations did, indeed, greatly stabilize the P_L transcript, and in these conditions, Nun caused a 2 to 4-fold decrease in the steady-state level of the *nut*-proximal transcript. In a second set of experiments, we inserted a DNA sequence encoding a stable RNA upstream of the *nut* site. We found that Nun decreased the steady-state level of this stable RNA 2 to 4-fold, in agreement with the results of the experiments carried out in the nuclease-deficient strain. These results suggest that Nun does, indeed, arrest transcription *in vivo*, but that eventually the arrested transcripts are removed from the template.

Proposed Course of Project:

1. Antitermination of early transcription in HK022. We plan to continue our characterization of the HK022 *nutL* site by constructing and analyzing mutants as described above. We will alter the predicted stem, terminal loop, internal loop, and bulge regions to determine their functional

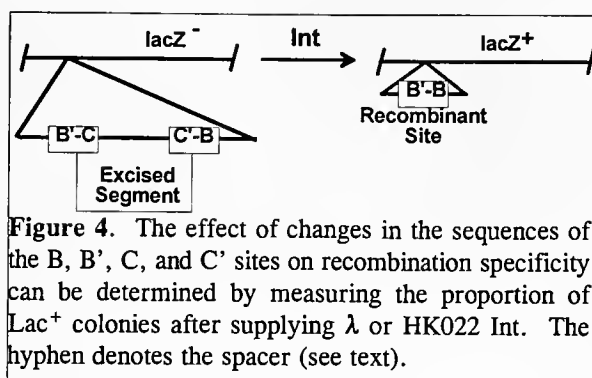


Figure 4. The effect of changes in the sequences of the B, B', C, and C' sites on recombination specificity can be determined by measuring the proportion of Lac⁺ colonies after supplying λ or HK022 Int. The hyphen denotes the spacer (see text).

importance. Compensatory mutations that restore base-pairing will provide evidence for the importance of a particular base-pair and against the importance of a particular sequence. Conversely, mutations that restore base-pairing but not function will suggest that sequence is important for function. We will try to select *nut* mutations that suppress other mutations that are located at positions that are predicted to be unpaired. We also plan to analyze the zinc binding region of the RNA polymerase β' subunit by alanine scanning mutagenesis. We will screen for haplo-lethal, dominant lethal, and antitermination-defective mutants. We will also ask if any mutations in this region can suppress antitermination-defective *nut* mutants, and, if so, if they mutation-specific or *nut*-independent suppressors. The same selective system is also suitable for isolating antitermination-resistant transcription terminators. If such terminators exist, we will analyze their function.

2. Antitermination with purified proteins. (i) We plan to test newly isolated *nut* and RNA polymerase mutants *in vitro* to confirm that antitermination with purified RNA polymerase faithfully reflects antitermination within cells. The efficiency of terminator readthrough *in vitro* is about one-half of what it is *in vivo*. We will attempt to increase the *in vitro* efficiency by adjusting ionic strength and the concentrations of Mg^{2+} and nucleoside triphosphates. If this does not succeed, we will search for host proteins that increase antitermination efficiency. (ii) Conversion of RNA polymerase to a persistently antiterminating state may change the kinetics of elongation including the propensity of polymerase to pause at pause sites. We plan to analyze the kinetics of transcription to see if the elongation rate or the likelihood of pausing is altered. (iii) Persistent antitermination is likely to involve a stable interaction of the *nut* transcript with polymerase. We will look for such an interaction in two ways. First, we will prepare synchronized ternary transcription complexes that are paused at particular positions on the template and look for RNA loops with the electron microscope. A stable interaction between *nut* and polymerase should produce RNA loops whose sizes depend on the position of the paused complex. Second, we will purify *nut* RNA and look for evidence of a interaction with RNA polymerase by interference with terminator readthrough, gel retardation and filter binding. If this experiment gives positive results, we will attempt to identify the region of polymerase that is bound to RNA by means of RNA-protein cross-linking and protein footprinting experiments.
3. Recognition of attachment sites by integrase proteins. (i) Binding of integrase to core binding sites is weak and difficult to detect under recombination conditions. However, if a DNA molecule that contains a core binding site also has an arm binding site, simultaneous binding of integrase to both sites promotes bending of the DNA, and such bending can be detected by binding of a second protein, such as HU, that has affinity for bent DNA. We plan to use this technique to ask if integrase mutations that switch or relax the specificity of recombination also alter binding to core sites of different specificities. (ii) We plan to construct and to characterize additional recombination substrates with core site mutations in order to learn which nucleotides of the cores sites are recognized differentially by the integrases of HK022 and λ . (iii) HK022 and λ integrases recognize identical core binding sites, and have identical core-binding domains. The two *int* genes are identical for the first (N-terminal) 154 bp, which encodes the core binding domain. The sequence identity continues, with a few interruptions, for several hundred base pairs upstream of the beginning of Int, but drops rather abruptly to approximately 70% identity downstream of position 154. It is likely that this transition point defines not only the beginning of an interdomain region, but also that of a crossover point between divergent ancestral *int* genes. If so, one of these ancestral *ints* has a core binding domain that differs from that of λ and HK022 *ints*. Evidence from genetic and molecular biological hybridization experiments suggests that such an *int* gene exists in the *E. Coli* chromosome. This gene could help us to define the amino acid residues that are responsible for core binding specificity, and we plan to clone and to characterize it.

4. The action of the Nun transcription-termination factor *in vivo*. We proposed above that the Nun transcription-termination protein arrests transcription *in vivo*, but that the arrested transcription complexes are removed from the template. Several cellular functions could potentially remove stalled transcription complexes: (1) The Rho transcription termination protein is believed to act on paused transcription complexes to cause termination; (2) The Mfd transcription-DNA repair coupling protein is believed to recognize transcription complexes that are stalled at lesions in DNA and to cause dissociation of these complexes from the template; (3) DNA replication could remove arrested transcription complexes. We are testing the first two hypotheses by determining the effect of *rho* and *mfd* mutations on the stability of *nut*-proximal transcripts in the presence of Nun. The third hypothesis will be tested by blocking DNA replication with thymine starvation, and measuring the effect on message stability.

Significance to Biomedical Research and the Program of the Institute:

The insertion and excision of viral and other chromosomes is known to change the physiology of the host organism. In addition, viruses provide an important mechanism of genetic transfer and exchange in the microbial world. Viruses promote genetic transfer because of their ability to convert host genes to an infectious form by packaging them inside a shell of viral structural proteins. Such a successful symbiotic relationship has a close analogy in the relationship between certain eukaryotic viruses and their hosts. An understanding of the control mechanisms operative in these easily studied microorganisms has obvious relevance to an understanding of regulatory elements in higher organisms.

Publications:

Clerget, M., Jin, D.J., and Weisberg, R.A.: A zinc binding region in the β' subunit of RNA polymerase is involved in antitermination of early transcription of phage HK022. *J. Mol. Biol.* **248**:768-780, 1995.

Dorgai, L., Yagil, E., and Weisberg, R.A.: Identifying determinants of recombination specificity: Construction and characterization of mutant bacteriophage integrases. *J. Mol. Biol.*, in press.

Gottesman, M. E. and Weisberg, R.A.: Termination and antitermination of transcription in temperate bacteriophages. *Seminars in Virology* **6**:35-42, 1995.

Nunes-Düby, S.E., Tirumalai, R.S., Dorgai, L., Yagil, E., Weisberg, R.A., and A. Landy, A.: Lambda integrase cleaves DNA in cis. *EMBO. J.* **13**:4421-4430, 1994.

Weisberg, R. A.: Specialized transduction. *Escherichia coli* and *Salmonella typhimurium*, Second Edition, in press.

Yagil, E., Dorgai, L., and Weisberg, R.A.: Identifying determinants of recombination specificity: Construction and characterization of chimeric bacteriophage integrases. *J. Mol. Biol.*, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

ZO1 HD 00067-27 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

The Integration of Macromolecular Synthesis in Escherichia coli

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	M. Cashel, MD PhD	Section Chief	LMG:NICHD
Others:	D. Gentry, PhD	IRTA	LMG:NICHD
	V.J. Hernandez, PhD	IRTA	LMG:NICHD
	D. Vinella	Visiting Fellow	LMG:NICHD
	H. Murphy	Chemist, Lab. Tech.	LMG:NICHD

COOPERATING UNITS (if any)

Dr. Gad Glaser, Hadassah Med. School, Jerusalem, Israel; Dr. Lilian Hsu, Holyoke College, Holyoke, MA; Dr. Arthur Kornberg, Stanford Univ. Med. Ctr. Stanford, CA; Dr. Horst Malke, Inst. Molec. Biol. Jena Univ., Jena GR

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Molecular Regulation

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

5.0

PROFESSIONAL:

4

OTHER:

1

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

This project involves understanding how changes in *Escherichia coli* gene expression are coordinated when cell growth is limited by nutrient availability. We are interested in signals provoking changes in cellular levels of regulatory nucleotide analogs of GDP and GTP that bear pyrophosphate residues on the ribose 3' hydroxyl; abbreviated as ppGpp and pppGpp, respectively, or (p)ppGpp collectively. Changes in (p)ppGpp levels are correlated with complex global changes in gene expression triggered by sensing nutritional depletion for sources of energy, of amino acids and of phosphate. Multiple sites for regulatory effects of (p)ppGpp probably co-exist at levels of metabolism, transcription and translation. One example in (p)ppGpp induction of expression of the alternative RNA polymerase sigma factor, RpoS, that triggers entry into a globally altered patterns of gene expression characteristic of stationary phase of growth. A 50X induction of RpoS protein occurs without appreciable change in rpoS mRNA levels. In another example, transcription is clearly affected. We reported that mutants of sigma-70, encoded by rpoD, suppress a feature of a (p)ppGpp-deficiency phenotype (multiple amino acid auxotrophy) that is a separate example of global positive regulation by (p)ppGpp. Here regulatory effects are exerted on transcript initiation, judging from altered patterns of abortive RNA chain release in vitro. Last year we noted mutant effects for model promoters with differing abortive RNA chain features; this year we found similar effects for the natural galP2 promoter. We now have verified that the sigma-70 mutants do not alter the kinetic constants for DNA binding or strand separation by RNA polymerase holoenzyme but again diminish RNA chain abortion in vitro while increasing the efficiency of full length transcripts clearly implicating promoter clearance. Resistance to mecillinam, a beta-lactam antibiotic, is a new selectable feature of elevated ppGpp levels; this year we have isolated mecR insertion mutants in relA-deleted hosts revealing several new possible sources of regulation of spoT functions.

Project Description

Objectives

We wish to understand how cellular gene expression is coordinated during adaptation to the stress of transitions between nutritional impoverishment and abundance. Nutritional stress is studied because it is arguably the most common biological stress. *Escherichia coli* is studied, with its full array of stress responses, because the molecular level of understanding of this organism is by far the most advanced of any organism capable of an independent existence. Although stress responses of other organisms will almost certainly show mechanistic differences, operating biological principles are likely to be shared. The genetic repertoire of *E. coli* is organized into functional domains or "regulons" that are operationally defined by shared regulatory responses to a given stress signal. Regulatory effects among various regulons can also be coordinated or linked, culminating in an integrated cellular response at the level of macromolecular synthesis and gene expression. Our overall objective is to understand how this works, starting with one global regulatory signal, (p)ppGpp, and then dissecting out regulatory interconnections.

The stress signals and macromolecules considered in this project have broadened considerably since the time it was originally aimed at the regulation of stable RNA accumulation mediated by the *relA* gene when amino acids or aminoacyl-tRNA availability limited exponential growth. The RelA protein is bound to ribosomes engaged in protein synthesis and responds to codon-specific binding of uncharged tRNA by activating (p)ppGpp synthesis. Energy source limitation also increases ppGpp abundance, but instead of stimulating (p)ppGpp synthesis, this stress leads to inhibition of (p)ppGpp breakdown. The major route of (p)ppGppase is encoded by the *spoT* gene, which encodes a (p)ppGpp 3'-pyrophosphohydrolase. Genetic evidence suggests that the SpoT protein is a bifunctional enzyme, catalyzing forward (synthesis) or reverse (degradation) reactions by a mechanism distinct from microreversibility in classical enzyme catalysis. When the *relA* gene is deleted, the sole remaining source of (p)ppGpp synthesis is judged to be the SpoT protein because when the *spoT* gene is also deleted, the cell is rendered devoid of detectable (p)ppGpp, abbreviated as (p)ppGpp⁰. Such strains, together with an engineered ability to gratuitously manipulate cellular levels of (p)ppGpp over a broad range has provided a general approach for learning the extent of participation of one global regulator in other responses. When these approaches are applied to cells in which genes that are judged to encode putative targets of (p)ppGpp are mutated, regulatory pathways can be mapped out. Thus, the objectives are: 1) to learn of the stress systems that provoke (p)ppGpp accumulation; 2) to understand the mechanisms involved in perturbations of (p)ppGpp metabolism; and 3) to define the regulatory effects on cell biology exerted by (p)ppGpp.

Methods Employed

Methods used have been published previously.

Major Findings

1. Characterization of RNA polymerase mutant sigma-70 subunit suppressors of a (p)ppGpp-deficient phenotype. Regulatory effects of (p)ppGpp have long been thought to occur at the transcriptional level but there is evidence favoring both initiation and elongation as the steps involved. In previous annual reports, isolation of mutants of RNA polymerase subunit genes (*rpoB*, *rpoC*, *rpoD*) were described as suppressors of several phenotypic features associated with a (p)ppGpp deficiency. We focused attention on the *rpoD* mutants because the steps affected in transcription were more likely to be limited and

identified two mutant sites (P504L and S506F) in a region between those responsible for recognizing the -35 and -10 promoter consensus elements. In vivo studies revealed that the mutants increased the fraction of Sigma-70 not in association with core RNAP but core-sigma association rates were not affected in vitro with purified components. Also last year, we reported that the reconstituted mutant enzyme affected abortive RNA chain formation using three sigma-70-dependent promoters (phage T7 A1, phage T5 N25 and "N25anti", a +3 to +20 sequence variant of N25). These promoters have been extensively characterized by Drs. L. Hsu and M. Chamberlin as sharing an unusual tendency to abort at high frequency but differing with respect to the chain length of abortively released RNA. The reiterative production of these RNA chains occurs without release of RNAP from the promoter. The process is halted when the enzyme embarks on processive elongation with the overall probability of promoter clearance apparently dependent upon sequences just downstream of the first few bases of the transcript. Using these promoters present on linear DNA, it was found that the mutant enzyme affected frequencies of short abortive RNAs in an interesting manner: 2-4 mers were unaffected, 5-8 mers were 3 to 10-fold reduced and increased for chains longer or equal to 9 bases in length. This RNA length- (or position-) dependent effect suggests that the mutant enzyme can "sense" differences while traversing the initially transcribed sequences and perhaps reflect different steps or stages in promoter clearance. Studies with these unusual promoters, while revealing, were qualified by two features. The first was a failure to observe a change in full length transcript abundance proportionate to changes in abortive RNA and the second was that ppGpp addition to wild type enzyme did not mimic mutant effects. We surmise that the mutant enzymes are not clearing these promoters with increased efficiency. They appear instead to alter the frequencies of abortive products while traversing the initial transcribed sequences without increasing promoter clearance.

We have now made extended our studies to include the galP1P2 tandem promoters from E. coli. These promoters were chosen because: 1) they have been well characterized (in other laboratories); 2) more than 99% of P2 initiations are known to be abortive; and 3) the ability to measure P2 specific abortive RNA products is greatly facilitated by a described technique for isolation of a supercoiled minicircle bearing only a single promoter. Minicircle galP1P2 plasmids were used as templates with wild-type and mutant holoenzyme under conditions where products from the P2 promoter could be assessed, despite the additional presence of the P1 promoter. It was found that the pattern and abundance of abortive RNAs was altered for the mutant enzymes but with some differences in chain length dependencies. The early abortive products (2-4 mers) were reduced in the mutants about 7-fold while the middle size products (5-6 mers) were 10- to 20-fold reduced and the longest abortive product, a 7 mer was only 2-fold reduced. In addition, the abundance of full length gal transcripts was increased collectively about 3- to 4-fold, implying that unlike the promoters mentioned above, the mutants were in fact decreasing abortive initiation and clearing the promoter at increased efficiency. To quantitatively localize the step responsible for the altered promoter activity, the physical constants K_b , associated with promoter binding, and k_f for the rate of open DNA complex formation were determined. No significant differences in these constants were observed between wild type and mutant enzymes with the galP2 promoter. Therefore it can be concluded that the sigma-70 mutants increase galP2 promoter strength by increasing the efficiency of promoter clearance. Transcription assays using minicircle templates in which only the galP2 promoter is activated gave identical patterns of abortive transcripts to those just mentioned, verifying that mutant effects on galP2 are indeed independent of galP1.

In vivo assays of effects of rpoD mutants on gal-lacZ fusion activities in (p)ppGpp⁰ hosts have revealed low levels of β -galactosidase with wild-type rpoD that are elevated by either the rpoD mutants or by restoring (p)ppGpp. This observation establishes a direct parallel between in vivo and in vitro behavior of the suppressor mutants in the absence of (p)ppGpp. However, addition of ppGpp to the galP2

transcription reactions with wild-type RNA polymerase did not alter full length product formation, an observation that does not coincide with in vivo behavior. Interestingly, additions of ppGpp to transcription reactions with mutant enzymes did further increase their promoter activity. Together this may be an indication that the mutations alter a requirement for an additional component normally mediating ppGpp-dependent effects. A mutant search for such a factor can be envisioned. Currently, it is uncertain whether the ability of the sigma-70 mutants to increase galP2 activity by increasing promoter clearance can be generalized to many promoters, to only promoters activated by ppGpp, or cannot be generalized. Whether the galP2 example serves as a special case or a model for positive ppGpp control of transcription awaits extension of these studies to other promoters.

Additional information is also anticipated from a variety of rpoB and rpoC mutants with suppressor activities analogous to the rpoD mutants just mentioned, and in many instances with substantially stronger activities. In addition, we do have additional rpoD mutants that are not localized to the region 3 class of mutants. We are in the process of working up these suppressor mutants to see if we can further define differences or similarities in mechanisms.

2. Mecillinam resistance as a positive selection for new mutants affecting ppGpp and targets of ppGpp regulation. Recently, the laboratory of Dr. Richard D'Ari (Paris) has found that mecillinam, a β -lactam antibiotic that specifically inactivates penicillin binding protein 2 (PBP2) results in cell death associated with inhibition of cell division. Cells become spherical in the presence of the antibiotic. This shape change is viewed as increasing the demand for FtsZ protein because viability is restored by overexpression of the ftsZ gene. The Fts protein is a tubulin-like protein whose polymerization is an early step in the formation of the septum between dividing cells. Studies in Paris have led to the isolation of two mecR mutants and identified them as lesions in aminoacyl-tRNA synthetase genes. Finding that the mecR phenotype of these mutants was reversed by relA mutants led to realizing that ppGpp accumulation was the source of the mecR phenotype. This was verified by gratuitous induction of ppGpp. These studies led D'Ari and colleagues to the hypothesis that ppGpp might function as a positive regulator of ftsZ and led us to realize that selections for the mecR in a relA deleted host could lead to the identification of new genes that can function in the regulation of elevated ppGpp levels by inducing spoT ppGpp synthesis, inhibiting spoT ppGppase or regulating ftsZ expression. The complex ftsZ operon has several promoters, some reportedly rpoS-dependent, which raises the possibility that ppGpp might indirectly induce ftsZ expression by inducing Sigma-S. However, this is apparently not the case; rpoS mutants are found incapable of reversing the mecR phenotype occasioned by ppGpp induction.

A search for mecR mutants in a relA1 host was initiated in Paris using random transpositions with a miniTn10::kan element. The two isolates obtained from this search were mapped in Paris as an insertion upstream of aroK in one case ("C3") and an unknown gene mapping at 99 min in the other ("B6"); both have been further characterized here this year. In addition, we have found 12 new insertions in a ΔrelA::kan host and a Tn10::cml minitransposon, six of these have now been localized by sequencing. Seven of the 12 become mecS in a ppGpp⁰ host, indicating ppGpp dependence, one shows intermediate sensitivity to mecillinam, and one is lethal. Three remain mecR in a ppGpp⁰ background, indicating ppGpp independence. Sequencing has revealed two insertions are in unknown genes now be sequenced. Identified insertion sites have been identified as follows: A10K = hns C-terminus; A10O = gef-ant intergenic region; A9B = pta orf; B10J = asnS early transcript; B10K = lpp N-terminus.

The insertion C3 (kan) is in the "early transcript region" between promoter for aroK and the translational start site for the encoded shikimate kinase. The aroK operon has the gene order: aroK aroB orf74.3 dam. A redundant shikimate kinase with a lower Km for shikimate is encoded by aroL

and it is necessary to mutate both aroL and aroK to obtain aromatic amino acid auxotrophs. The two shikimate kinases are not equivalent with respect to mecR since aroL::Tn10 remains mecS. The aroK-specific mecR phenotype cannot be explained by a polar effect of the aroK insertion on downstream operon genes since aroB, orf74.3 and dam mutants remain mecS and since a plasmid borne minimal aroK gene (made by PCR) complements the mecR phenotype of the aroK insertion. The mecR phenotype is also not due to perturbations of the aromatic amino acid biosynthetic pathway since a complete block before (aroB) or after (aroA) the shikimate kinase step does not generate mecR and does not complement the C3 insertion. The AroK protein is likely to have two functions. An aroK C-terminal region missense mutant (on a plasmid) has been isolated that complements the mecR C3 insertion phenotype but fails to complement the shikimate kinase deficiency of an aroK aroL double mutant. This seems to eliminate shikimate kinase activity itself and raise the possibility of involvement of a different phosphorylation event. The mecR feature of C3 persists in relA1 ΔspoT but not in ΔrelA ΔspoT strains, indicating that it is ppGpp-dependent.

The insertion B6 (kan) is in an unknown gene and its genetic instability prevents usual tests for ppGpp dependence. This insertion was found in Paris to display a 70-100% increased FtsZ content, suggesting that the insertion might have inactivated a negative regulator of FtsZ expression.

The insertion A10D (cml) confers intermediate mecR in a ppGpp⁰ strain and the insertion site is also in an unknown gene.

The insertion A10K (cml) has been mapped to 27 min and identified as an insertion in the C-terminus of hns coding for a DNA binding protein with known pleiotropic effects. A Δhns mutant was found to similarly display a ppGpp-dependent mecR phenotype. Mutants in hns have been found to decrease rmBP1-lacZ fusion activity (contrary to predictions in the literature) specifically in minimal medium. The slow growth of a Δhns is suppressed by a rpoS mutant (other laboratories) and found to be exacerbated by a purR disruption, which we believe may sensitize cell growth to ppGpp. We currently deduce that hns lesions either elevate ppGpp levels or contribute to sensitizing cell growth to ppGpp.

The insertion A100 (cml) confers ppGpp-independent mecR and has been localized to the intergenic region between divergent gef and ant genes. The former gene is implicated in an addiction system while the latter affects Na⁺/H⁺ antiporter activity. We have not yet studied which of these activities are related to mecR.

The insertion A9B (cml) confers ppGpp-dependent mecR and affects the pta gene, responsible for acetyl phosphate and acetyl-CoA interconversions. We are now testing effects of deletions of pta as well as other genes encoding enzymes involved in the metabolism these compounds. Interestingly, acetyl phosphate regulation has been implicated in regulation of phosphorylation of PhoB, a DNA binding protein regulator of expression of the pho regulon, another global regulatory circuit. It is also intriguing that phosphate deprivation is known to lead to PhoB phosphorylation, to pho regulon derepression and to ppGpp accumulation in a spoT-dependent manner.

The insertion B10J (cml) occurs in the "early transcript" region of asnS, an essential gene coding for asparaginyl-tRNA synthetase. Interestingly, the insertion is lethal in a ppGpp⁰ host. The mecR phenotype of the B10J insertion persists in ΔrelA hosts, unlike other aminoacyl-tRNA synthetase mutants with a mecR phenotype. The basis for this behavior is unknown.

The insertion B10K (cml) confers a ppGpp-dependent mecR phenotype and is localized in the N-terminus of the lpp gene coding for a major lipoprotein. This gene has been well studied by others but we have not applied this information to explore the significance of this finding with regards to mecR.

We have also tested effects on the mecR phenotype of sequenced rpoB and rpoD M⁺ suppressors of the multiple amino acid auxotrophy of ppGpp⁰ strains. Such an effect might be expected because

they represent mutations that mimic ppGpp positive control and because an unsequenced rpoB mutant allele was shown in the D'Ari laboratory to block cell division at high temperature and abolish the mecR phenotype normally produced by moderate elevation of ppGpp; both features were suppressed by further increases of ppGpp or by a plasmid overproducing the FtsZ protein. The two rpoD alleles mentioned above confer mecR in a ppGpp⁰ strain. Of the 10 rif-r M⁺ suppressor alleles, 4 gave a mecR phenotype in both wild-type and ppGpp⁰ hosts; 2 were mecR in the wild-type but not the ppGpp⁰ strain, and 4 were mecS in both backgrounds. All of the M- rifampicin-resistant alleles gave a mecS phenotype. The ability of these RNA polymerase subunit mutants to give a mecR phenotype was well correlated with an independent ranking of the hierarchy of suppressor strengths. This evidence underscores the conclusion that mecR probably reflects a positive control function of ppGpp.

3. (p)ppGpp induction of Sigma-S formation. We reported earlier that (p)ppGpp-dependent induction of Sigma-S accompanied limiting exponential growth for lack of available amino acids, energy sources, or phosphate. Furthermore, Sigma S contributes to changes in gene expression accompanying these growth limiting conditions. We suspect that the importance of this response is that (p)ppGpp induction from a variety of sources of stress acts as a premonitory signal of impending stationary phase that occurs before severe nutritional exhaustion actually occurs.

The mechanism by which (p)ppGpp gives a 50X induction of rpoS gene expression within 10 min remains elusive. It cannot be quantitatively explained by mRNA abundance during induction although levels of rpoS mRNA do not increase with cell density in a ppGpp⁰ strain as they do in a wild type strain. If rpoS mRNA levels are manipulated to increase in a ppGpp⁰ strain, translation does not increase proportionately. We have verified reports that RpoS protein stabilization occurs late in stationary phase and found this effect is (p)ppGpp-independent; however only minor changes in protein stability occur during ppGpp induction. The translatability of rpoS mRNA isolated from (p)ppGpp-induced cells in S30 bacterial cell extracts is not found to be appreciably altered, suggesting there is not a mRNA modification change facilitating translatability. The possibility exists that translational efficiency of an apparently constant amount of rpoS mRNA is enhanced indirectly by (p)ppGpp induction of degradation of other mRNA species. Attempts to mimic this effect with rifampicin addition have revealed indeed that Sigma-S levels do increase when its mRNA is declining; the turnover rate of rpoS mRNA is 3.5-4 min. A similar effect on Sigma-S abundance is obtained when mRNA turnover is provoked by pyrimidine starvation of an auxotroph but inexplicably not by purine starvation. Interestingly, another laboratory has argued that the ability of ppGpp induction to reduce translational errors during amino acid starvation might be due to alleviation of the demand for charged tRNA by global reduction of mRNA levels. It should be stressed that such a reduction has been deduced, and not yet documented.

If mRNA levels are indeed reduced by the presence of ppGpp, how might this occur? The recent literature has reports from Kushner's and Cohen's groups that polyA tails, synthesized by polyA polymerase encoded by pcnB, can activate mRNA degradation by facilitating binding of RnaseE-polynucleotide phosphorylase complexes. It has long been known that the copy number of plasmids with a variety of replication origins is reduced by ppGpp accumulation; the mechanism is unknown but it is intriguing that the pcnB mutation was isolated and the gene named "plasmid copy number" for its negative effects on plasmid copy number, now appreciated to operate by influencing polyA tail length to affect the competition between sense RNA and antisense RNA that determines priming ability. It seems plausible that ppGpp might affect polyA polymerase activity and thereby explain

both its effects on plasmid copy number and its presumed effects on global mRNA levels. These speculations can be tested with existing pcnB deletion mutants.

Levels of Sigma-S increase about 20X in 15 min after osmotic shock (.3 M salt) but only about 3X in a ppGpp⁰ strain; osmotic shock has long been known to provoke ppGpp accumulation in a spoT gene-dependent manner. Over this same period, the level of rpoS mRNA increases about 2X. Other labs have reported similar osmotic shock effects on rpoS mRNA and Sigma-S abundance and also suggested a posttranscriptional mechanism. We suspect that the mechanism of induction is not unique to osmotic shock but instead is another manifestation of ppGpp induction. The significance of osmotic shock as an inducer of ppGpp-mediated Sigma-S expression is that it is a stress that presumably does not involve nutritional deprivation.

At this point regulatory effects of ppGpp on rpoS gene expression seem complex. Effects seem unlikely to occur at the level of initiation of transcription; ppGpp does not appreciably induce any of the several promoters in the rpoS operon but the absence of ppGpp weakly reduces all, despite very different promoter sequences. Levels of Sigma-S rise where overall transcription is inhibited (by induction of ppGpp, rifampicin or pyrimidine starvation) and repressed when transcription is stimulated in a ppGpp⁰ background. Translation efficiency of rpoS mRNA could be passively regulated by competition with other mRNA species although this is difficult to define because estimates of bulk mRNA (as well as individual mRNA) turnover rates vary depending on the technique used. Indeed, our estimates of rpoS mRNA stability after rifampicin treatment could be complicated by translation affecting degradation rates. Nevertheless, increased Sigma-S expression in the presence of rifampicin seems to be unique; we are unaware of similar reports for any other proteins. Alternatives to passive regulation of rpoS mRNA translational efficiency would be the direct effects of ppGpp on translation or the existence of an repressor of rpoS mRNA translation that is negatively regulated by ppGpp. The first alternative can be tested with an S30 translation system and the second alternative can be explored by searching for a putative translational repressor gene using a rpoS'-lacZ protein fusion with the fusion junction near the C terminus.

4. SpoT protein activities for ppGpp synthesis and degradation. The spoT gene encodes a complex source of (p)ppGpp regulation because it functions as both a (p)ppGpp synthetase and a (p)ppGppase, which would be a futile cycle if unregulated. We have now further localized the domains responsible for each activity and find both in the N-terminal half of the protein with a central overlapping area and have mutants that show exclusively one, the other, or neither activity. These are being exploited to study regulation of each function in isolation of the other. In multicopy the isolated synthetic function appears to be an unregulated activity. This assessment is being made under the more meaningful condition of low copy, where putative regulatory factors are not titrated out by overexpression. Plasmids with a spoT gene fragment encoding ppGpp degradation activity but not synthetic activity has been used in a Δ spoT relA⁺ host to reassess contributions of RelA to carbon source starvation. We find in such strains that when dextrose is limited in the presence of amino acids, no ppGpp accumulates; when amino acids are absent, a transient burst of ppGpp appears. This suggests that the RelA system does doesn't respond to glucose deprivation but that this limitation does gives a transient amino acid deficiency, which RelA does respond to. We have also devised a dual plasmid system in which the balance between synthesis and degradation of (p)ppGpp may be manipulated to select unique functional mutants in one or the other plasmid. Since regulation of (p)ppGpp is spoT gene-dependent for virtually all stresses other than amino acid limitation, the mechanism of regulation of these functions will soon become a central issue.

Publications:

Cashel M, Gentry D, Hernandez, VJ, Vinella D. The stringent response. In: Escherichia coli and Salmonella typhimurium cellular and molecular biology. (Ed. F. C. Neidhardt), Americ. Soc. Microbiol. Press, Washington, D.C. 1995 in press.

Gentry DR, Cashel M. Cellular Localization of the Escherichia coli SpoT protein. J. Bacteriol. 1995;177:3890-3893.

Hernandez VJ, Cashel M Changes in conserved region 3 of Escherichia coli σ^{70} mediate ppGpp-dependent functions in vivo. J. Molec. Biol. 1995; in press

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 00068-24 LMG

PERIOD COVERED

October 1, 1994 through September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Factors Influencing Genetics Transcription-Initiation and Termination

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	Robert J. Crouch, Ph.D.	Senior Investigator	LMG:NICHD
Others:	Wu-Po Ma	Visiting Fellow	LMG:NICHD
	Susana Cerritelli	Senior Staff Fellow	LMG:NICHD
	Zhan, Xinyi	IRTA	LMG:NICHD
	Jonathan Bennett	Pre-IRTA	LMG:NICHD

COOPERATING UNITS (if any)

Section on Viral Gene Reg., LMG:NICHD; Protein Expression Lab., OD:NICHD; Dr. Valarie Mizrahi, Mol. Biol Unit, The S. African Inst. for Med. Res., Johannesburg; Dr. Marc Drolet, Dept. of Microbiol & Immunol, Univ of Montreal, Canada

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Formation of RNA

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

4.7

PROFESSIONAL:

4.7

OTHER:

0.0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Ribonuclease H (RNase H) plays crucial roles in replication of retroviruses, including the human immunodeficiency virus (HIV) as well as involvement in normal cellular events including DNA replication. The purpose of this project is to gain a detailed understanding of the precise nature of the substrates for RNases H, the chemical makeup of cellular RNases H. The enzymes derived from cellular sources have higher specific activities than those of retroviral origin and have somewhat different specificities. We have found that separation of the RNase H domain of MuLV reverse transcriptase from the polymerase domain creates an active enzyme with specificities different from the parental RNase H but still with low specific activity. If we produce RNases H from yeasts, they have high specific activities regardless of their association with their own special domains. These results indicate that the differences in these proteins reside in the RNase H domain itself.

1800

1810

1820

1830

1840

1850

1860

1870

1880

1890

1900

1910

1920

1930

1940

1950

1960

1970

1980

1990

2000

2010

2020

2030

2040

2050

2060

2070

2080

2090

2100

2110

2120

2130

2140

2150

Project Description:Objectives:

The aim of this project is to examine the involvement of certain ribonucleases in cellular events. The scope of this project has evolved to include protein chemistry, enzymatic activity, overproduction of proteins and the consequences of changes in the level of enzymatic activity within the cell. In particular, ribonucleases specific for RNA of RNA-DNA hybrids are being examined for their functional role within the cell. Multiple enzymes exhibiting RNase H activity have been reported within cells and our experiments are aimed at understanding the need for multiple forms of the enzyme and what effects there would be upon inhibition of each (or all) of the RNase H activities. This is particularly relevant since the AIDS virus requires the functioning of its own RNase H to be infectious.

Major Findings:

Proteins that have RNase H activity are found in bacteria, plants, animals and in viruses such as HIV. While all of these enzymes can recognize and degrade the RNA of RNA-DNA hybrids, they are diverse in (i) primary amino acid sequence (only 9 of 155 amino acids are conserved between the HIV RNase H and that of *E. coli*) (ii) enzymatic activity (both in preference for divalent metal ions and in specific activity - the range being greater than 10^5) yet appear to have similar tertiary structures. The lack of sequence homology amongst the various RNase H genes may important for the cell to maintain multiple genes that have RNase H activity.

I. Conserved motif in eukaryotic RNases H and a transcriptional activator**a) Amino acid sequences**

Two copies of a sequence in *S. cerevisiae* RNase H1, a single copy of a similar sequence in *S. pombe* RNase H1 and in an RNase H from *Crithidia fasciculata*, a trypanosome, has recently been found in a sequence deposited in the EST database from chicken. Last year, we reported that this region seems to be responsible for binding to double-stranded RNA and RNA-DNA hybrids. We have tried to define the amino acids involved in dsRNA binding and have reached the following conclusions: (i) The first copy of the dsRNA-binding motif is necessary and sufficient for strong binding to dsRNA. (ii) The domain can be divided into two regions based on very high conservation in the first part(A) and less conservation in the second part(B). (B has the most similarity to other known dsRNA-binding sequences). Replacement of B of the *S. cerevisiae* RNase with B from either *C. fasciculata* RNase H or from the cauliflower mosaic virus Transactivator protein eliminates the dsRNA-binding, suggesting that parts A and B may require non-conserved amino acids for proper interaction with dsRNA.

b) Size of dsRNA bound

Originally, we used a 4 kilobase pair dsRNA for binding studies. Most dsRNA-binding proteins require about 20-30 base pairs for binding and for full activation, PKR kinase requires greater than 30 base pairs. To gain more insight into the minimal length of dsRNA required for binding, we have tested a variety of smaller dsRNAs for interactions with *S. cerevisiae* RNase H1 on the 39 kDa (full length) RNase H1 and on the N-176 protein comprised of the amino terminal domain (the dsRNA-binding region). From results reported last year, we know that the RNase H domain and the dsRNA-binding domain influence each other. For example, the 39 kDa protein binding to dsRNA is almost completely inhibited at 5 mM Mg^{2+} whereas the N-176 protein binding is reduced only about half that of the maximum observed at low or no Mg^{2+} . N-176 seems to bind equally to the 4-kbp dsRNA and a 37 bp dsRNA. dsRNAs of 8, 11 and 20 bp also bind to N-176 (with equal efficiency) but their binding is

much weaker than the 4 kbp and 37 bp RNAs. To our surprise, several of the mutant versions of the N-176 protein fail to bind the 4 kbp dsRNA but exhibit some interaction with the 8 bp dsRNA. The TAT protein of HIV binds to an RNA in HIV RNA called TAR, a hair-pin containing about 100 nucleotides. N-176 and 39 kDa proteins bind to the TAR sequence but not as efficiently as to an 84 bp dsRNA. The binding assay used for these experiments is a Northwestern assay in which the various proteins are blotted to filter paper and then incubated with radiolabeled dsRNAs. We have started using the BIAcore for obtaining binding constants and we are seeing strong binding of the 39 kDa protein to a 36 bp dsRNA but much weaker binding to the 8 bp dsRNA.

II. Differences between cellular and retroviral RNases H

We have reported the three-dimensional structure of *E. coli* RNase H1. Others subsequently showed that the HIV reverse transcriptase (RT) RNase H domain is very similar in overall structure to that of the *E. coli* enzyme.

Last year, we reported that the isolated RNase H domain of HIV RT can be converted from an inactive to an active form by substituting the "handle" region of HIV with the corresponding region from the *E. coli* enzyme. Interestingly, the HIV-*E.coli* chimeric enzyme cleaved one test substrate at the same site as the complete RT enzyme. *E. coli* RNase H1 has a specific activity more than 1,000 times that of retroviral RNases H. The chimeric RNase H specific activity resembled the RT enzyme, not that of the *E. coli* protein. To understand what elements of the two enzymes contribute to the differences in specificity and specific activity, we have extended our studies to include the RNase H of AKR murine leukemia virus (MuLV). Unlike the HIV enzyme, the AKR-MuLV RNase H domain expressed by itself has RNase H activity equivalent to that when it is attached to the polymerase domain. It also retains the strong preference for Mn^{2+} over Mg^{2+} characteristic of the MuLV RT. However, when using poly(rA)-poly(dT) as substrate, the purified RNase H domain exhibits a cleavage pattern markedly different from that of the RNase H activity of MuLV reverse transcriptase. The differences in cleavage were further revealed by using tRNA^{pro} model substrate. While the MuLV RT cleaves the RNA-DNA junction on this defined substrate, the purified RNase H domain cleaves this substrate only at other sites. This result is in contrast to the results reported last year with the HIV-*E. coli* chimeric protein. In addition, degradation products observed using two other substrates were different depending on whether the RT or the RNase H domain of AKR-MuLV was used. Replacement of the "handle" region of MuLV with its *E. coli* homologue did not alter the specificity or specific activity. These results suggest that the RNase H domain of MuLV reverse transcriptase retains a full RNase H activity, it loses cleavage specificity. Examination of the products generated by the HIV-*E. coli* chimeric enzyme and the HIV RT with these additional substrates showed that there are differences observed, depending on the substrate used. These results are consistent with the conclusion that AKR-MuLV and HIV RNases H differ in their requirements for the polymerase domain for activity but both require the complete RT for accurate cleavage.

Publications:

Journal Articles

Cerritelli SM, Crouch RJ. The non-RNase H domain of *Saccharomyces cerevisiae* RNase H1 binds double-stranded RNA: Magnesium modulates the switch between double-stranded RNA binding and RNase H activity, RNA 1995;1, 246-59.

Dawes SS, Crouch RJ, Morris SL, Mizrahi V. Molecular cloning, sequence analysis, overproduction in *Escherichia coli* and enzymatic characterization of the ribonuclease HI from *Mycobacterium smegmatis*, Gene, in press.

Drolet M, Phoenix P, Menzel R, Masse E, Liu LF, Crouch RJ. A major role for DNA topoisomerase I in *E. coli*: inhibition of R-loop formation during transcription elongation, Proc Natl Acad Sci U.S.A. 1995;92, 3526-30.

Guo J, Wu W, Yuan ZY, Post K, Crouch RJ, Levin JG. Defects in primer-template binding, processive DNA synthesis and reverse transcriptase activity associated with chimeric reverse transcriptases having the Murine Leukemia Virus polymerase domain joined to *Escherichia coli* RNase H, Biochemistry 1995; 34, 5018-29.

Stahl SJ, Kaufman JD, Vikić-Topić S, Crouch RJ, Wingfield PT. Construction of an enzymatically active ribonuclease H domain of human immunodeficiency virus type 1 reverse transcriptase, Protein Engineering 1994;7, 1103-08.

Journal Supplement

Crouch RJ, Tozawa Y, Cerritelli SM. Ribonucleases H of *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and *Crithidia fasciculata* share a common motif in their non-RNase H domain that binds double-stranded RNA in a magnesium-dependent manner. Protein Engineering 1994;7, 1151-2.

Future Work:

I. Targeting RNase H to retroviral-like particles of *S. cerevisiae* and to HIV-1 virions

We have been attempting to ask some basic questions related to replication of retroviral-like elements in *S. cerevisiae*. These TY-1 elements replicate using reverse transcriptase and require RNase H in this process. We have asked what happens when we express *E. coli* RNase H in *S. cerevisiae* in a manner in which it gets incorporated into virus-like particles. What we have found is that transposition can be inhibited 100% if we use a wild-type RNase H gene but no inhibition is observed when an enzymatically inactive version of RNase H is introduced. We are currently trying to assess the exact step in transposition inhibited by the *E. coli* RNase H. However, the question arises as to whether these results can be obtained in the HIV-1 system. To carry out experiments in the more complex human system, we are expressing *E. coli* RNase H fused to the gag protein of HIV-1 by using plasmid vectors that can replicate in *E. coli* and mammalian cells. Each plasmid will be introduced into HeLa-CD4+ cells and the affect of expressing the fusion protein on HIV-1 infection will be assessed. If we get results suggesting that *E. coli* RNase H can also interfere with HIV-1 replication, we will attempt to determine the substrate that the RNase H recognizes by looking at various replication intermediates. Such inhibition would be useful both as a tool for understanding replication of HIV-1 and, in theory, could be used for gene therapy of patients with the HIV-1 virus. We will also, in collaboration with Jef Boeke, Johns Hopkins University, measure the effect of expressing a murine leukemia virus gag-*E. coli* RNase H fusion protein.

The *S. cerevisiae* TY-1 system can be exploited to examine the effects of various retroviral reverse transcriptase chimeric enzymes that have been or will be constructed. It has been suggested that it may

be possible to replace the TY reverse transcriptase with that of MuLV and maintain transposition. If this turns out to be true, we can rapidly screen a number of chimeric proteins for their effects on "viral" replication. Most of the mutants we have now are defective for replication in vitro (see report by J. Levin, LMG, NICHD). Expression of these mutant RTs in the TY system could also be used to select for RTs that are active in vivo. This approach is asking for the RT to function when a foreign or altered RNase H is present and should be contrasted with the selection for "better" RNase H activity described below.

II. Selection of active RNase H genes by complementation of *E. coli* temperature-sensitive strain

We have used our *E. coli* strain to select for genes that encode RNases H. This system can also be utilized to select for "better" RNase H genes. For example, the E-loop (*E. coli* - HIV chimeric RNase H) functions poorly in the presence of Mg^{2+} but has increased activity in the presence of Mn^{2+} . Can we get complementation with this construct, provided we include $MnCl_2$ in the plate and will we be able to select for mutants that can complement without $MnCl_2$. This strategy can be applied to other constructs that fail to complement for a variety of reasons (e.g., unstable protein, miss-folded protein, etc.). We now have several chimeric RNases H consisting of portions of the RNase H region of *E. coli* and part of the AKR MuLV RNase H domain. We have tested all of these for complementation but none actually does. By selecting for clones that do complement and determining the changes, we expect to gain more insight into what actually makes a good RNase H.

III. Search for active RNase H genes

A) Modification of PCR strategy: We attempted, unsuccessfully, to obtain RNase H genes from higher eukaryotic cells using the PCR method that yielded the *S. pombe* *RNH1* gene. The current methodology relies on a set of degenerate oligonucleotides for amplification. As the genome of organisms examined become more complex, there is a much greater chance that the "correct" oligonucleotide primers will be consumed very early in the PCR reaction. To circumvent the depletion of the relevant oligonucleotides, we will synthesize degenerate oligonucleotides that have unique sequences at their 5'-ends as well. After a few rounds of PCR amplifications under stringent conditions, we will add oligonucleotide primers that correspond to these unique sequences and continue to amplify the PCR products.

B) Search for dsRNA-binding sequence homologies and screening of libraries for dsRNA-binding proteins: Characterization of the *RNH1* protein of *S. cerevisiae* has given us a new (but limited) handle on a specific type of RNase H - one that binds dsRNA. We will develop a screen of expression libraries in which we use labelled dsRNA as probe. Since the *RNH1* protein binds to dsRNA in the absence but not in the presence of Mg^{2+} , we can search for those dsRNA-binding proteins that have this unusual property. An alternative method (which may be employed) would be to purify proteins from other organisms that bind to dsRNA columns and can be eluted by addition of $MgCl_2$. We are currently testing to see if this procedure is effective for purification of the *RNH1* protein. If we can obtain such proteins, microsequencing of the protein should give us enough information for making PCR primers.

C) DataBase searches: The accelerated rate of genomic and cDNA (EST) sequences being deposited in the data bases makes searching by computer a useful approach. We have several RNase H gene sequences with which to search. This is important since the most convenient tool for searching is the BLAST program. In our searches, this program is very sensitive to having the conserved amino acids separated from each other by a defined number of non-conserved amino acids. We are currently

studying two genes that may encode RNase H from mouse and from human. These sequences can be obtained and cDNA libraries can be screened to obtain the complete gene or cDNA.

IV. Chimeric *E. coli*-retroviral RNase H proteins

We have constructed a series of chimeric RNases H between *E. coli* and AKR MuLV proteins with the goal of understanding the differences in specific activities and divalent metal ion requirements of these two enzymes. With this system, we will be able to construct RTs that have modifications in their RNase H domain for study of the interaction of the RNase H and polymerase domains. Each of these chimeric proteins will be candidates for use in "Targeting RNase H in *S. cerevisiae*" - see I. above. We are also in the process of making a HIV RNase H that has a single cysteine residue which will be used to link a deoxyoligonucleotide which may be specific for cleavage of RNA complementary to the short DNA. This will confirm that the HIV RNase H is defective in binding to nucleic acid - not in its catalytic activity.

V. Yeast RNases H1

A) Cell localization: One of the key questions about the RNH1 protein of *S. cerevisiae* is what is the cell location of this enzyme? We have tried several approaches to determine where the enzyme resides. In *S. cerevisiae* the protein seems to be present in very low amounts. Our best antibodies do not detect the enzyme on Western Blots even though low levels of the protein are easily detected in *E. coli* carrying the yeast gene. We are constructing two genes that will express the yeast RNase H fused to two reporter genes that may help in cell localization. One reporter gene is the green fluorescent protein (GFP). This reporter protein can be detected in living cells and its presence in different cellular compartments can be followed during cell growth, at different stages of the cell cycle and under various conditions. The second reporter gene is one in which the yeast gene is expressed with a His-tag permitting rapid isolation of the protein from cell extracts. This latter construct may also be useful for determining what type of nucleic acid is bound to the enzyme (see below). For the *S. pombe* RNase H we are using the β -galactosidase reporter gene. This reporter protein can be easily assayed in crude extracts and will help in measuring the levels of RNase H (see below).

B) dsRNA-binding motif:

i) Related proteins: *S. pombe*, *C. fasciculata* and chick RNase H all have regions very closely related to the dsRNA-binding motif of *S. cerevisiae* RNase H1. A similar sequence is also present in the translational transactivator of cauliflower mosaic virus. We are examining the ability of each of these to bind dsRNA and are asking if we can replace portions of the non-conserved regions with amino acids found in the other. With this, we hope to understand the amino acids required for binding.

ii) Size and sequence specificity: We are examining a series of dsRNAs ranging in size from 12 to 36 base pairs for their ability to bind to various dsRNA-binding proteins. One oligo in each pair has a biotin label and can be linked to a chip on the BIAcore plasmon resonance machine. Using this instrument, we can obtain binding constants and learn the minimum size of dsRNA to bind. The sequence of the oligos has no biological significance. To search for the *in vivo* substrate, we can use a protein bound to BIAcore chip and screen for RNAs that bind best to the protein. Fractionation of RNAs into various size classes or RNAs isolated from cytoplasmic or nuclear fractions can be used to begin this type experiment. It will also be possible to use the SELEX procedure for synthesis and selection of RNAs that have tight binding. This latter technique will be useful when we can determine a "minimum size" dsRNA (as described above). Another approach to discovering the *in vivo* substrate

employs the His-tagged RNase H described in Section III. Part A. Isolation of the protein from yeast cells using a nickel affinity column may yield a protein to which its normal substrate remains bound. If we can obtain enough of this material, it may be possible to identify the nucleic acid using a variety of techniques, including PCR.

VI. Differences between cellular and viral RNases H

Results presented under Major Findings indicate that specific activity differences between cellular and retroviral RNases H reside in the RNase H domain itself and is not due to totally to an association with the polymerase domain. If one compares the amino acid sequences of all RNases H there are about 9 amino acids scattered throughout the protein that are conserved. Two sets of sequences are found to be conserved in cellular enzymes but not viral proteins. We plan to modify the viral proteins to include these sequences to see if they play a role in specific activity levels.

VII. Cellular perturbations and RNase H levels

The presence of a dsRNA-binding domain on cellular RNases H raises the possibility that there is a connection between turning off - or down - translation of mRNAs by PKR in response to dsRNAs and RNase H regulation by dsRNA. To address this question, we will study the levels of RNase H under different conditions known to increase, decrease or activate PKR. One such example is a change in PKR during apoptosis. We will follow the changes in RNases H activity in murine cell lines undergoing this programmed cell death.

77
R
T
P
S
V
T
C
R
D
P
H

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 00069-23 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Molecular Genetics of Mammalian Retrovirus Replication

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

P.I.:	J.G. Levin	Research Biochemist	LMG:NICHD
Others:	J. Guo	IRTA Fellow	LMG:NICHD
	M. Powell	IRTA FELLOW	LMG:NICHD
	W. Wu	Visiting Fellow	LMG:NICHD
	Y.-X. Feng	Guest Worker	LMG:NICHD
	K. Post	Biologist	LMG:NICHD

COOPERATING UNITS (if any)

ABL Basic Research Program, NCI-FCRDC (A. Rein, T.D. Copeland); SAIC Frederick, NCI-FCRDC (L. Henderson); NICHD-LMG (R. Crouch)

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Viral Gene Regulation

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

5.0

PROFESSIONAL:

4.0

OTHER:

1.0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The goal of this project is to define the molecular mechanisms involved in the replication of mammalian retroviruses and in particular, to understand the factors which influence the regulated expression of viral genetic information. We have been using an in vitro approach to study events in reverse transcription. We have shown that during (-) strand DNA elongation on a murine leukemia virus (MuLV) RNA template, there is a major pause site preceding a stable stem-loop structure, formed by bases from the polypurine tract (PPT) and downstream sequences. HIV-1 nucleocapsid (NC) protein, which destabilizes secondary structure, reduces pausing and increases the rate of synthesis of full-length DNA. This shows that NC stimulates processivity of reverse transcriptase (RT). Experiments with chemically modified NC reveal that the cysteine residues in the Zn fingers are required for NC activity in the pausing assay and facilitate NC binding to the template. In studies with HIV-1 RT, we have found that initiation of (+) strand viral DNA synthesis is dictated by the PPT alone and not by the surrounding context. Based on the results of assays measuring the ability of wild-type and mutant PPT RNA oligonucleotides to act as primers and CD spectra of short PPT-containing RNA-DNA hybrids, we also conclude that a combination of an unusual helical structure and a requirement for certain bases at the 3' end of the PPT appear to be important determinants for primer recognition and extension. Analysis of RNase H* activity (degradation of double-stranded (ds)RNA) shows that an active RNase H domain is required; activity is enhanced by fusion of the RNase H domain to polymerase, which stimulates binding of the dsRNA substrate to RT. In studies on in vitro formation of retroviral RNA dimers, we have demonstrated that NC protein can convert an unstable dimer to a more thermostable form. Thus, we now have an in vitro model for NC-mediated maturation of the genomic RNA dimer in virus particles. Experiments to test the activity of mutant NC proteins in the pausing and dimerization assays are in progress.

Project Description:

Objectives: To define in molecular terms the mechanisms involved in the replication of mammalian retroviruses, in particular, murine leukemia virus (MuLV) and human immunodeficiency virus (HIV); to study factors which influence the regulation and expression of viral genetic information; to understand the relationship between genetic structure and functional activity in virus infection.

Methods Employed: Recombinant DNA technology is used for molecular cloning of viral genes. DNA fragments to be cloned are amplified by the polymerase chain reaction (PCR) procedure. Cloned viral DNAs are expressed in *E. coli* or in some experiments, are transferred to mammalian cells by standard transfection techniques. Viral and plasmid DNAs are analyzed by restriction endonuclease digestion followed by agarose gel electrophoresis. DNA sequence analysis is carried out by the dideoxy-chain termination method. Viral DNA sequences cloned in expression vectors are transcribed with either SP6 or T7 RNA polymerase. The T7 transcripts are used for analysis of RNase H activity and as templates in primer extension assays. In some cases, SP6 and T7 transcripts from the same DNA plasmid are hybridized to form a substrate for assay of a retroviral dsRNA nuclease activity. Bacterially expressed reverse transcriptase (RT) is purified by chromatography on HPLC and Sephacryl S-200 columns. Viral proteins are analyzed by immunoprecipitation or by immunoblotting techniques. SP6 transcripts are used for studies on in vitro dimerization of viral RNA. Dimer formation is analyzed by electrophoresis in 2% Metaphor gels followed by ethidium bromide staining; for labeled RNAs, phosphorimager analysis is used.

Major Findings:1. Reverse transcription

A. Processive DNA synthesis. Studies have been initiated to investigate factors affecting processivity of reverse transcriptase (RT). This question is clearly of great interest since the efficiency of viral DNA synthesis will have a direct effect on the efficiency of virus replication. In earlier work we noted that during (-) strand elongation on an MuLV RNA template, there is a major pause site at a position preceding the polypurine tract (PPT). Pausing occurs just before a run of 4 C's and results in formation of a 75-nt DNA. Examination of the kinetics of DNA synthesis showed that over a 1 h time interval, the amount of full-length product (321 nt) is increased, while that of the 75-nt DNA is decreased. This result is consistent with a precursor-product relationship between the two DNAs. Secondary structure prediction with the RNA-fold program suggests that a stable 24-nt stem-loop structure can be formed by interaction of some bases in the PPT, including the 6 G's, and bases downstream of the PPT, including the 4 C's. If either the 6 G's or 4 C's are changed to a random sequence, then the putative stem-loop is not expected to form and pausing should not occur. However, if a compensatory mutation is made (switch 4 C's and 6 G's), the program predicts a new, but similar stem-loop with comparable stability, which would impede DNA synthesis. These predictions were realized in primer extension assays with the mutant templates and strongly supported the idea that a stem-loop structure near the MuLV PPT is responsible for the observed RT pausing. Since nucleocapsid (NC) protein can destabilize secondary structure, we also tested the effect of HIV-1 NCp7 in the primer extension assay. NCp7 reduced pausing near the PPT and increased the rate of synthesis of full-length DNA. This shows that NC stimulates RT processivity.

HIV-1 NCp7 has two Zn fingers, each with a highly conserved Cys-His motif (C-X₂-C-X₄-H-X₄-C); the Zn fingers are required for viral RNA packaging and infectivity. To assess the influence of the Cys



residues in the Zn fingers on RT pausing, we have used a drug, N-ethylmaleimide (NEM), which specifically targets these residues. We have found that chemical modification of NCp7 with 3 equivalents of NEM, which binds exclusively to the 3 Cys in the second Zn finger, partially reduces NC activity; modification with 6 NEM completely abolishes activity. Results of a filter binding assay demonstrate that NCp7 with 3 NEM is impaired in its ability to bind a T7 viral RNA transcript, whereas NCp7 with 6 NEM binds little or no RNA. These findings indicate that the Cys residues in both Zn fingers play a role in the ability of NC to reduce pausing, presumably by facilitating binding of NC to the template. (Collaboration with L. Henderson and T.D. Copeland, FCRDC)

B. Requirements for Initiation of Plus-Strand Viral DNA Synthesis. Selective RNase H cleavage at the 3'-terminus of the PPT generates the RNA primer for (+) strand DNA synthesis. Last year we reported that a conserved U-rich region flanking the 5'-end of the PPT is not essential for (+) strand priming. We have now found that moving the 15-nt PPT or the PPT plus the first 5 downstream bases to two different contexts does not affect priming activity. From these results, we conclude that initiation of (+) strand DNA synthesis is dictated by the PPT alone and not by interactions with sequences in the immediate or more distant 3' and 5' contexts.

To investigate elements in the PPT responsible for specific primer recognition and extension by RT, we have been testing the activity of short RNA oligonucleotides annealed to longer ss-DNAs in a primer extension assay. We find that the 15-nt PPT can be efficiently extended by HIV-1 RT and T4 polymerase, but a 15-nt downstream sequence is only active with the T4 enzyme. Mutations in the PPT are also being analyzed. The HIV-1 PPT sequence is 5'-AAAAGAAAAGGGGGG; the position of the first G at the 3' end is designated -1, the second G, -2 etc. To date, the results show that mutations at the 5' end of the PPT do not affect priming activity, whereas 3' mutations do have an effect. Indeed, alteration of the 4 G's in positions -1 to -4 to 4 C's completely abolishes activity. Surprisingly, a mutant PPT which retains the 6 G's at the 3' end, but has a random sequence in the other 9 nt, is fully active in the primer extension assay. We are also using CD spectroscopy to probe the helical structure of PPT-containing RNA-DNA hybrids. Wild-type PPT and mutant PPTs that can be extended have a characteristic CD spectrum not shared by inactive variants, indicating that the biologically active and inactive PPT sequences are in distinct structural classes. Interestingly, an oligonucleotide (5'-GAGAGAGAGAGAGAG) which is in the same class as the PPT based on structural considerations and CD spectrum, has some priming activity, but cannot be extended as efficiently as the wild-type PPT.

We conclude that (i) the PPT alone is sufficient to initiate (+) strand DNA synthesis; and (ii) a combination of an unusual helical structure and requirement for certain bases at the 3' end of the PPT appear to be important determinants for primer selection.

C. RNase H* Activity. The ability of retroviral RT to degrade dsRNA is termed RNase H* activity. In the course of investigating this activity, we unexpectedly found that *E. coli* RNase H, generally used as a negative control, has a low level of RNase H* activity. Fusion of *E. coli* RNase H to the MuLV polymerase domain (resulting in a chimeric RT, as described previously) increases RNase H* activity by ~10-fold. Maximal stimulation of RNase H* activity is obtained with the entire polymerase domain. Thus, the activity of a chimeric RT missing most of the connection subdomain is only ~3-fold greater than that of *E. coli* RNase H. Not surprisingly, a mutation in one of the RNase H catalytic residues (D10A) abolishes the RNase H and RNase H* activities of *E. coli* RNase H and the chimeric RT. Filter binding experiments support the conclusion that the polymerase domain enhances RNase H* activity by stimulating binding of the dsRNA substrate to RT. This work represents an additional

example showing that one domain in RT contributes to the activity of the other.

2. Dimerization of Retroviral RNA

Last year we reported that three distinct regions within a larger 342-nt Harvey sarcoma virus RNA can dimerize in vitro. Two of the smaller RNAs form dimers with a melting profile like that of virion RNA, while the third forms unstable dimers. Interestingly, dimers of the larger 342-nt RNA exist in two forms: dimers formed at 55°C are more thermostable and melt at a higher temperature than dimers formed at 37°C. This finding is reminiscent of the maturation of dimeric RNA in virions, in which there is a conformational change converting an “immature” dimer to a more stable form. Because NC has been implicated in RNA dimer maturation in virus particles, we have been investigating the effect of HIV-1 NCp7 on in vitro dimers formed at 37°C. We find that the melting profile of the 37°C dimer treated with NC is indistinguishable from that of untreated 55°C dimer. These results demonstrate that NC can stabilize the structure of a retroviral RNA dimer and provide the first in vitro model for maturation of dimeric viral RNA. Experiments with mutant NC proteins are in progress. (Collaboration with A. Rein, L. Henderson, and T.D. Copeland, FCRDC).

3. Immune Response to HIV-1 and HIV-2 Peptide Antisera

Previously we reported Western blot analysis showing that HIV-2 peptide sera made against sequences in RT or integrase (IN) cross-react with HIV-1 RT and IN proteins, respectively. Cross-reactivity of sera from HIV-2 positive individuals was also noted. In contrast, HIV-1 peptide and patient sera are type-specific. Recently, we have also observed a functional cross-reactivity. HIV-2 antibodies made against the N-terminus of HIV-2 IN inhibit HIV-1 IN-catalyzed cleavage of a DNA LTR substrate. Taken together, the results on HIV-2 cross-reactivity have added significance in view of a very recent clinical study carried out in Senegal showing that infection of individuals with HIV-2 protects against infection with HIV-1. This suggests that development of an HIV-2 based vaccine may provide a forceful strategy for AIDS prevention.

Proposed Course of Project:

1. Reverse Transcription

A. Role of NC Protein: (1) Processivity. In experiments described above, we have shown that HIV-1 NCp7 stimulates RT processivity on an MuLV RNA template and reduces pausing near the PPT. Moreover, results obtained with NCp7 modified with NEM indicate that the cysteine residues within the Zn fingers are important for NC activity in this assay. In a more detailed analysis, we are now testing the activity of mutant NC proteins to identify residues in NC which are required to reduce RT pausing. Dose response, kinetic experiments, and filter binding assays are being performed. Initial results demonstrate that the eight basic amino acids flanking the first Zn finger are essential, since mutation of these residues to Ala completely abolishes NC activity. Other mutations to be analyzed include: point mutations in the Cys and aromatic residues in the Zn fingers; deletion of the entire first or second Zn finger; and rearrangement of the Zn fingers. We will also test the activity of NC proteins from other retroviruses, such as SIV (simian immunodeficiency virus), EIAV (equine infectious anemia virus), HTLV-1, and BLV (bovine leukemia virus). [Note: MuLV NC, which has only one Zn finger, seems to be unstable compared with HIV-1 NCp7; to date we have been unable to detect activity with either synthetic MuLV NC or NC isolated from virions.] (2) Strand Transfer (First Template Switch). Following synthesis of (-) strong-stop DNA at the 5' end of viral RNA, there is a template switch, and

the DNA is repositioned at the 3' end of the genome for further elongation. We plan to set up an in vitro assay for strand transfer, first with sequences from the 5' and 3' LTR sequences of HIV-1. Both donor and acceptor templates will be synthesized as T7 transcripts and a DNA primer will be used; DNA products will be analyzed on a 6% sequencing gel. It has been reported that NC protein stimulates strand transfer in vitro. We will perform experiments to confirm this finding using HIV-1 NCp7 in our system. We will also test the activity of mutant NCp7 proteins and heterologous retroviral NC proteins (see above). If heterologous NC turns out to have the ability to stimulate strand transfer, we will set up an MuLV strand transfer system using HIV-1 NCp7 and exploit this system to analyze mutations in the polymerase domain of MuLV RT.

B. Determinants of MuLV RT Function. (1) Active form of MuLV RT. Last year we reported that the RNase H domain of MuLV RT facilitates formation of a super-shifted complex in a band-shift assay with RT and a short DNA primer-template; formation of this complex is favored when binding is followed by primer extension. We considered the idea that this complex may represent the active form of RT, possibly a dimer. We are now interested in determining which regions within the RNase H domain and elsewhere in RT are required for primer extension and formation of the super-shifted complex. One approach we are using is to test the inhibitory effect of synthetic peptides on these activities. In this way we hope to identify residues which are involved in generating the active form of RT. In a survey of six MuLV peptides from the RNase H domain, we have found that only one inhibits primer extension; not surprisingly, this peptide contains sequences from the region required for substrate binding. Since the HIV-1 connection subdomain has been implicated in formation of the p66/p51 RT heterodimer, we have also been testing peptides (30-35 amino acids) from the MuLV RT connection subdomain. To date, we have found two overlapping peptides which significantly inhibit primer extension. We plan to have additional peptides synthesized to ascertain which residues are critical for the observed inhibition. We will also conduct parallel tests to see which peptides have inhibitory activity in the band-shift assay. Using a different approach, in collaboration with Dr. R. Fisher, FCRDC, we are exploring the possibility of studying the kinetics of binding of MuLV RT to primer-template with or without concomitant DNA synthesis. This work would involve the use of the BIAcore machine; the data would be evaluated to determine whether they are consistent with formation of an RT dimer. Initial experiments are being performed with HIV-1 RT, which is known to form a heterodimer and can serve as a model for MuLV experiments. (2) Mutations in the MuLV Polymerase Domain. Using cassette mutagenesis, we plan to make a series of mutations in the polymerase domain of MuLV RT. In the first of these experiments, we will make point mutations in conserved residues in a region analogous to the "primer grip" of HIV-1 RT. We will measure activity in RNase H, primer extension, and strand-transfer assays.

2. Requirements for Initiation of Plus-Strand DNA Synthesis. (1) Until now, much of the work with the short RNA primers has been performed with an assay procedure in which single-stranded phagemid DNA serves as the (-) strand DNA template and product analysis depends on primer extension to define the position where DNA synthesis is initiated. We have now modified this assay as follows: (i) a 35-nt DNA oligonucleotide, whose 3' end is complementary to the 5' end of the RNA primer, is used as the template; and (ii) DNA products are labeled directly by inclusion of [³²P]dATP in the reaction. This assay is more convenient and allows us to compare the relative activities of different primers in a more quantitative manner. Experiments are now in progress to test wild-type and mutant PPTs for specific cleavage at the 3' boundary of the PPT, primer extension, and primer removal. To further define sequence determinants at the 3' end of the PPT, we will also test the activity of additional mutant RNA primers. (2) In addition to the in vitro work, we plan to investigate sequence requirements for initiation of (+) strand DNA synthesis in vivo, using a transient transfection assay. This work will be a

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

collaboration with Dr. A. Rein, FCRDC. We will clone the Harvey sarcoma virus (HaSV) genome into a mammalian expression vector pcDNA3, which has a CMV promoter. HaSV has a PPT which is very similar to that of AKR or Moloney MuLV (it is 14 nt rather than 13 nt long, with the added base near the 5' end); the genome also contains sequences required for packaging viral RNA and sequences encoding *ras*, but not *gag*, *pol*, or *env* proteins. The HaSV clone will be transfected into a packaging cell line which synthesizes all of the retroviral proteins required for virus assembly. To assay for focus formation by the HaSV-containing particles, the supernatant fluids from transfected cells will be collected after 24-48 hours and then used to infect NIH3T3 cells; foci should be detected after several days. To specifically probe for the activity of the PPT, we will use PCR analysis to detect (+) strand DNA in the infected cells, at various times post-infection. Parallel experiments will be carried out to test the activity of HaSV wild-type and mutant PPT sequences in an in vitro system, like the one described above for HIV-1.

3. In Vitro Formation of Retroviral RNA Dimers. Experiments are now in progress to test the ability of mutant HIV-1 NCp7 proteins (see above) to convert the unstable 342-nt HaSV RNA dimers formed at 37°C to the more thermostable form. We are interested in comparing NC function in the pausing assay, which appears to depend on NC binding to the RNA template, with NC activity in the dimerization assay, where NC also catalyzes a conformational change in the RNA. Preliminary data suggest that the requirement for the cysteine residues in the Zn fingers may differ in the two assays. The variety and number of Zn finger mutants provided by our collaborators should allow us to resolve this point. Experiments are also planned with NC proteins from other retroviruses (see above).

Significance to Biomedical Research and the Program of the Institute:

The advantages of viral systems as models for studying the regulation and expression of genetic information are well recognized. Studies on retroviruses are of special interest since these viruses establish a chronic infection in the host cell by integrating viral genetic information into the host chromosome, in some instances disrupting normal gene function. An understanding of the various steps involved in virus replication is crucial for the success of efforts to limit the pathogenic potential of these agents. More specific considerations follow. (i) Among molecular biologists there is wide interest in the process of reverse transcription. Evidence from non-viral systems indicates that processed genes, certain highly repeated short mammalian DNA segments, and genes for small nuclear RNAs may have been generated by such a mechanism. It has also been known for some time that many eukaryotic transposable elements have a structural organization remarkably similar to that of the retroviruses. The discovery that yeast Ty elements encode a functional reverse transcriptase and that an RNA intermediate is involved in transposition of the element, has lent strong support to the idea that transfer of information from RNA to DNA has been an important factor in the development of the eukaryotic genome. It is interesting that reverse transcriptase is also found in certain bacteria. (ii) The clinical importance of retroviruses should also be emphasized. Retroviral vectors have been considered for some time for use in somatic gene therapy and a number of clinical trials with such vectors are now underway. In addition, public attention continues to be urgently focused on the AIDS epidemic. Studies on the MuLV system have been invaluable for elucidating the molecular mechanisms involved in the replication of the human retrovirus, HIV, the etiological agent responsible for AIDS. At present, reverse transcriptase is the only target of anti-HIV drug therapy for infected individuals.

Publications:

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

Feng YX, Fu W, Winter AJ, Levin JG, Rein A. Multiple regions of Harvey sarcoma virus RNA can dimerize in vitro, J Virol 1995;60:2486-90.

Guo J, Wu W, Yuan ZY, Post K, Crouch RJ, Levin JG. Defects in primer-template binding, processive DNA synthesis, and RNase H activity associated with chimeric reverse transcriptases having the murine leukemia virus polymerase domain joined to Escherichia coli RNase H, Biochemistry 1995;34:5018-29.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01002-13 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Gene Expression During Embryonic Development of *Xenopus laevis*

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	I.B. Dawid	Head	(All Personnel Listed LMG:NICHD)	
Others:	A. Karavanov	Vis. Sci.	M. Rebagliati	IRTA Fellow
	M. Taira	Vis. Sci.	R. Toyama	Vis. Associate
	M. Rebbert	Chemist	M. Kobayashi	Guest Researcher
	J-P.Saint-Jeannet	Vis. Fellow	P. Curtiss	Biologist
	E. Glasgow	NRC Fellow	W. Hane	Guest Researcher
	S. Rjazansky	Guest Res.	T. Tanaka	Guest Researcher
	T. Tomita	Guest Res.	B. Huang	Summer Student

COOPERATING UNITS (if any)

H. Westphal, LMGD, NICHD; R. Harland, U.C. Berkeley; K. Ozato, LMGR, NICHD; K. Cho, UC Irvine; Xi He & H. Varmus, NIH; Judith Eisen, U. Oregon

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Developmental Biology

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

9.6

PROFESSIONAL:

7

OTHER:

2.6

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Molecular mechanisms regulating embryogenesis of the amphibian *Xenopus laevis* and the zebrafish have been studied. Pattern formation is controlled in the embryo by the action of cell signalling molecules that elicit the expression of transcription factor-encoding genes in a spatially and temporally regulated way. The focus of this laboratory has been the study of the earliest establishment of dorsoventral polarity and the function of LIM class of homeobox genes whose members are instrumental in pattern formation and tissue differentiation.

The dorsoventral axis is established in the *Xenopus* embryo in very early development through a signalling pathway that can be initiated by certain Wnt factors. We have explored the signal transduction cascade of these factors by studying GSK-3, a kinase believed to be involved in the process. By expression of wild type and dominant negative GSK-3 constructs it was shown that inhibition of GSK-3 activity dorsalizes while activation of GSK-3 ventralizes the embryo. These reciprocal effects strongly support the involvement of GSK-3 and of the entire Wnt signalling pathway in dorsal axis formation.

The *Xlim-1* gene is expressed in the organizer region of the embryo. Expression of mutant forms of *Xlim-1* can turn animal (ectodermal) cells into organizer-like cells in that they secrete factors that induce neural tissue in ectoderm and muscle in ventral type mesoderm. One of the factors whose production is elicited by activated *Xlim-1* is the recently discovered signalling protein chordin; chordin may be a major signalling factor elaborated by the organizer. The nature of the induced neural tissue is anterior, as seen by the expression of several marker genes. The transcription factor *Xbra* acts to convert the anterior neural tissue induced by *Xlim-1* into more posterior tissue without itself being a neuralizing agent.

One of the key issue in studying early development is the transcriptional control of activation and maintenance of regulatory genes that are important at this time. Both the regulation of transcription of *Xlim-1* and its activation of a target gene have been studied. A major form of regulation of gene activity in development is mediated by growth factors, and *Xlim-1* is activated by the TGF- β -related factor activin. This activation has been shown to depend on a constitutive promoter in the upstream region of the gene together with an activin-responsive silencer located in the first intron. The *goosecoid* gene is a target for *Xlim-1* regulation. A region in the *goosecoid* upstream region has been shown to be required for a response to *Xlim-1*, and binding sites for the *Xlim-1* homeodomain have been identified in this region. This work begins to establish a network of regulatory interrelationships during the gastrula period of embryogenesis in a vertebrate embryo.

Project Description:Objectives:

This project aims to provide insights into molecular mechanisms during early animal development. *Xenopus laevis* has been chosen as the main experimental system because it is a vertebrate with easily accessible and well-studied embryos. Much information is already available on the *Xenopus* developmental genes, the eggs and embryos of this animal are suitable for a wide variety of experimental manipulations, and many molecular markers have been isolated that are specific for different developmental stages and different regions of the embryo. Thus, considerable background information exists in this system. However, standard genetics cannot be studied effectively in *Xenopus*, and therefore a research program has been initiated in the zebrafish where such studies are possible. A major aspect of these studies concerns the establishment of the body plan and initiation of tissue differentiation, which in vertebrate embryos depends to a very large extent on inductive cell interactions. A major consequence of cell interactions is the activation of regulatory genes in a spatially and temporally specific manner; such events have been studied in both *Xenopus* and the zebrafish. The focus has been on the LIM class of homeobox genes which includes several members that are involved in pattern formation and tissue differentiation.

Methods Employed:

Many recombinant DNA methods are being applied, including cDNA and genomic DNA cloning, hybridization to RNA and DNA on filters and in solution, restriction site mapping and DNA sequencing, primer extension analysis, polymerase chain reaction (PCR), and S1 nuclease/RNase protection assays. Bacterial expression of the protein products of some of the studied genes has been used to investigate some properties of the proteins and to generate antibody probes. In vitro transcription and translation is used for the characterization of proteins encoded by isolated genes. Cytological methods include in situ hybridization to tissue sections and in whole mounts, and immunochemical staining of tissue sections and whole mounts. Classical embryological manipulation techniques, such as microdissection, cell disaggregation and explant culture are being utilized. Injection of cloned genes and constructs derived from them into embryos is used to study transcriptional control of the genes. Injection of synthetic RNA into oocytes or embryos is used to express the corresponding proteins in vivo for an analysis of their properties and their functions.

Major Findings:

The LIM subclass of homeobox genes in *Xenopus*. As reported in previous years, we have characterized three LIM class homeobox genes from *Xenopus* and the zebrafish and, in collaboration with H. Westphal and colleagues, the mouse. The expression patterns of these genes implicate them in multiple developmental processes, notably the formation and patterning of the central nervous system (CNS). These studies have been extended and completed in the past year.

We have focused on the function of *Xlim-1* because of its expression in the Spemann organizer region of the early gastrula. To study *Xlim-1* function, RNA transcribed in vitro from this gene was injected into frog embryos and the consequences of overexpression of the *Xlim-1* protein were studied. These experiments have suggested that LIM domains are negative regulatory domains. While wild type protein is inactive, *Xlim-1* mutant protein is capable of several functions assigned to the Spemann organizer. These functions include the ability to induce the neural markers in ectoderm, as well as

markers for the cement gland, a non-neural anterior structure. In cooperation with the mesoderm-specific Xbra product, Xlim-1 can also induce muscle-specific actin. These effects were shown to be indirect, requiring additional cell communication. While the putative inducing factors noggin and follistatin are not involved, the recently discovered factor chordin is induced by mutant Xlim-1 protein; chordin may be a major transducer of inductive influences in the gastrula. The neural tissue induced by Xlim-1 has an anterior character, as is the case for all primary induced neural tissue. The question which agents posteriorize parts of the neural plate, thus organizing the CNS, is an important one. We found that the transcription factor Xbra is capable of generating a signal that converts anterior to posterior neural tissue; however, Xbra is not itself a neural inducer.

The Xlim-1 protein activates expression of the homeobox gene *gooseoid* (*gsc*), which is also expressed in the Spemann organizer of the gastrula. In collaboration with Ken Cho, the regulation of *gsc* by Xlim-1 has been studied, showing that Xlim-1 activation depends on 430 nts of *gsc* upstream region. Several binding sites for the Xlim-1 homeodomain have been identified in this region.

Xlim-1 is regulated by activin in animal explants. Using appropriate constructs, it was possible to reconstruct this regulation in injection experiments. An upstream region of about 300 bp from the *Xlim-1* gene acts as a constitutively active promoter; its function is suppressed by sequences in the first intron. Constructs containing the first intron are inactive unless treated with activin, implying the existence of an activin-sensitive silencer in this region of the *Xlim-1* gene. The region required for this activin response has been localized to a 210-nt long segment approximately in the middle of intron 1. The *Xenopus Xlim-1* constructs are also regulated when injected into zebrafish embryos; both activin, and the distantly related growth factor nodal, activate the construct in the heterologous embryo. These experiments highlight the conservation of regulatory pathways during vertebrate evolution.

GSK-3 is involved in signalling in the early embryo. GSK-3 is a kinase with several known in vitro substrates that is of developmental interest because its *Drosophila* homolog acts downstream of signals mediated by the *wingless* and *Notch* genes. In collaboration with Xi He and Harold Varmus we have shown that inhibition of GSK-3 by way of dominant negative constructs leads to the generation of a dorsal axis; in contrast, overexpression of wild type GSK-3 ventralizes the embryo. Further, wild type GSK-3 antagonizes the effect of Wnt-8. These results indicate that (i) a Wnt signal transduction pathway is active in the early amphibian embryo and both necessary and sufficient for dorsal axis formation; (ii) GSK-3 is part of the Wnt signalling cascade in vertebrates; (iii) the GSK-3-mediated signal arises early, and is upstream of the ventralizing signal mediated by BMP-4 since BMP-4 overcomes the dorsalizing effect of GSK-3 inhibition.

Retinoic acid regulates development in frogs and fish in different ways

In collaboration with K. Ozato and others we tested the developmental effects of several retinoids in *Xenopus* and zebrafish embryos. It is well known that all-trans retinoic acid (t-RA) leads to anterior and posterior malformations in frog and other vertebrate embryos. Retinoid effects are mediated by two types of receptors, RARs and RXRs; the former are believed to require ligand binding for activity, the latter may also bind ligands but act as heterodimer partners. Thus it was surprising to find that RXR-specific ligands had very similar biological effects in *Xenopus* as RAR-specific or non-specific ligands. In zebrafish embryos, the situation is quite different: general and RAR-specific ligands act as teratogens, but RXR-specific ligands have no effect. These results suggest the possibility that additional and as yet unrecognized receptors exist in *Xenopus* that can mediate the effect of the novel ligands.

Neural Induction in *Xenopus*

Neural induction has been studied in collaboration with Richard Harland, Berkeley. Several marker genes isolated in this laboratory and specific for different regions of the nervous system, were tested for inducibility by the novel inducing factor noggin. We found that the different markers responded differentially to the inducing factor. The most significant aspect of this study is the observation that naive ectoderm, when induced by noggin, can self-organize into anterior-posterior and dorsal-ventral patterns. This observation has interesting implications for the patterning in the early CNS.

Significance to Biomedical Research and the Program of the Institute:

Molecular events that occur before and during gastrulation are of critical importance to the entire developmental program of any vertebrate organisms. This project takes advantage of distinct features of the amphibian and fish embryos to study these events in some detail. In particular, it has been clear for a century that inductive interactions are key events for the establishment of body pattern and tissue differentiation in the embryos of all animals and especially in those of vertebrates. Mesoderm induction is the earliest known cell interaction in the vertebrate embryo, followed closely by the elaboration of the nervous system. These events determine the formation of the body axis and the differentiation of major tissues. Therefore a thorough analysis of the molecular mechanisms of mesoderm induction and early neurogenesis is an important part of the overall aim to understand the basis of normal and abnormal embryogenesis. The analysis of regulatory genes that are active during this critical stage of development is a particularly promising approach to this overall aim.

Proposed Course:

The studies proposed for the future include a continuation of the molecular analysis of early embryogenesis in amphibian and fish. In particular, we propose to extend the analysis of the LIM class of genes in several ways. The inactivity of the wild type *Xlim-1* protein as compared to mutants that modify or remove the LIM domains suggests that this protein must be activated in vivo to carry out its functions. We believe that LIM proteins are activated in vivo by association with different co-activating proteins. Therefore, a search for LIM-domain binding proteins is underway. One such protein has been isolated in the Westphal Laboratory and its properties are being studied in collaboration. Preliminary evidence suggests that this protein specifically binds LIM domains but not mutated LIM domains, and that the binding protein can interact with *Xlim-1* in vivo to activate biological functions. These aspects will be pursued in both a molecular and biological direction. Additional LIM-domain binding proteins will be sought with the aid of the yeast two hybrid system.

The mechanism of activation of the *gsc* gene by *Xlim-1* will be pursued. The regions required for *Xlim-1* binding will be further delineated, and the possible cooperation of *Xlim-1* with other factors will be tested.

The regulation of expression of the *Xlim-1* gene will be studied further. We plan to localize the activin response element more precisely. Next, we will identify the transcription factor(s) that bind to this region and determine whether these factors are responsible for the in vivo activation of *Xlim-1* at the midblastula transition. We also plan to investigate these effects in the heterologous system of the zebrafish. Further, we shall test whether the activin responsive element also responds to other factors, e.g., nodal-related factors, Wnt factors, and fibroblast growth factor.

In the zebrafish we propose to characterize the nodal-related genes that have recently been isolated in our laboratory. These genes may be involved in pattern formation at the gastrula stage, and are thus of considerable interest. After basic gene characterization we will prepare suitable constructs for embryo injection to test their biological activity. In collaboration with the Oregon zebrafish group the map positions for all new fish genes cloned in this laboratory will be determined; we already have map positions for some of these genes. This opens the promising approach to compare the location of the various cloned genes with that of mutations generated in the zebrafish in several laboratories. A new direction is being initiated in the zebrafish. Using γ -rays we will generate large chromosomal mutations (deletions, rearrangements) in zebrafish sperm. After mating to wild type females, the deletion of interesting genes can be tested by PCR of individual haploid embryos. In this way we propose to look for mutants in the *lim1*, *lim3*, *lim5* and *lim6* genes. If such mutations can be obtained, they will provide useful material for phenotypic analysis and for the screening of point mutations in these genes.

In collaboration with Xi He we continue to analyze the role of GSK-3 in embryogenesis. The role of the GSK-3 pathway in neural induction and in the Notch pathway will be studied.

Publications:

Brown DD, Dawid IB. Introduction. *Curr Opin Genet Dev* 1994;4:627-9.

Burgener-Kairuz P, Cortesy-Theulaz I, Merillat AM, Good P, Geering K, Rossier BC. Polyadenylation of Na(+)-K(+)-ATPase beta 1-subunit during early development of *Xenopus laevis*. *Am J Physiol* 1994;266:C157-64.

Dawid IB, Toyama R, Taira M. LIM domain proteins. *Compt Rend Acad Sci Paris* 1995;318:295-306.

Good PJ. A conserved family of *elav*-like genes in vertebrates. *Proc Natl Acad Sci USA* 1995;92:4557-61.

He X, Saint-Jeannet JP, Woodgett JR, Varmus HE, Dawid IB. Glycogen synthase kinase 3 and dorsoventral patterning in *Xenopus* embryos. *Nature* 1995;374:617-22.

Jungwirth C, Rebbert M, Degen HJ, Schultz U, Dawid IB. Chicken ICSBP and IRF-1 genes reveal evolutionary conservation in the IRF gene family. *Proc Natl Acad Sci USA* 1995;92:3105-9.

Knecht AK, Good P, Dawid IB, Harland RM. Dorsal-ventral patterning and differentiation of noggin-induced neural tissue in the absence of mesoderm. *Development* 1995;121:1927-36.

Saint-Jeannet JP, Karavanov AA, Dawid IB. Expression of mesoderm markers in *Xenopus laevis* Keller explants. *Int J Dev Biol* 1994;38:605-11.

Saint-Jeannet JP, Dawid IB. A fate map of the 32-cell stage of *Rana pipiens*. *Dev Biol* 1994;166:755-62.

Taira M, Otani H, Saint-Jeannet JP, Dawid IB. Role of the LIM class homeodomain protein Xlim-1 in neural induction and muscle induction by the Spemann organizer in *Xenopus*. *Nature* 1994;372:677-9.

Toyama R, Curtiss P, Otani H, Kimura M, Dawid IB, Taira M. The LIM class homeobox gene *lim5*: implied role in CNS patterning in *Xenopus* and Zebrafish. Dev Biol, in press.

Toyama R, O'Connell ML, Wright CVE, Kuehn MR, Dawid IB. Nodal induces ectopic *gooseoid* and *lim1* expression and axis duplication in zebrafish. Development 1995;121:383-91.

Xu RH, Kim J, Taira M, Srendi D, Kung HF. A dominant negative bone morphogenetic protein 4 receptor causes neuralization in *Xenopus* ectoderm. Biochem Biophys Res Comm 1995;212:212-9.

Zhadanov AB, Bertuzzi S, Taira M, Dawid IB, Westphal H. Expression pattern of the murine LIM class homeobox gene *Lhx3* in subsets of neural and neuroendocrine tissues. Dev Dynamics 1995;202:354-64.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01004-12 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Regulation of Amino Acid and Nucleotide Biosynthesis in *Saccharomyces cerevisiae*

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: A.G. Hinnebusch, Research Microbiologist, (All listed personnel LMG:NICHD)
M. Garcia Vis. Fellow M. Marton IRTA Fellow G. Pavitt Visiting Fellow
C. Vasquez Vis. Fellow P. Romano IRTA Fellow J. Anderson IRTA Fellow
W. Yang Vis. Assoc. R. Rolfe IRTA Fellow E. Duenas Guest Researcher
B. Jackson Biologist K. Natarajan Vis. Fellow
C. Drysdale IRTA Fellow F. Zhang (Ebon)
L. Phan Guest Researcher

COOPERATING UNITS (if any)

None

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Molecular Genetics of Lower Eukaryotes

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

11.06

PROFESSIONAL:

10.06

OTHER:

1.0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

We are studying transcriptional and translational control mechanisms that regulate amino acid and purine biosynthetic genes in yeast in response to nutrient availability. GCN4 is a transcriptional activator of amino acid biosynthetic genes that is regulated at the translational level by the eIF-2 α kinase GCN2. We have proposed that phosphorylation of eIF-2 by GCN2 under starvation conditions stimulates GCN4 translation by inhibiting guanine nucleotide exchange on eIF-2 catalyzed by eIF-2B, thereby reducing the concentration of the ternary complex eIF-2/GTP/Met-tRNA^{Met}. A large number of point mutations were isolated in the GCN3, GCD7, and GCD2 subunits of eIF-2B that uncouple GCN4 translation from eIF-2 α phosphorylation without affecting eIF-2B catalytic activity. These mutations cluster in two regions of sequence similarity shared among the three proteins which may define an interaction surface between eIF-2B and eIF-2(α P). Many of the mutations affect residues conserved between yeast and mammalian eIF-2B, and the rat homologue of GCN3 can partially substitute for GCN3 in yeast cells. GCD2, GCD7, and GCN3 can form an eIF-2B subcomplex in vivo that reverses the effects of eIF-2 α phosphorylation, probably by sequestering eIF-2(α P). We have shown that GCD10 is a component of the 8-subunit eIF-3 complex, and that mutations in GCD10 are suppressed by overexpressing tRNA^{Met}. These results suggest that *gcd10* mutations reduce binding of ternary complexes to 40S ribosomal subunits. GCN1 and GCN20 are components of a protein complex required in vivo for phosphorylation of eIF-2 α by GCN2. An N-terminal segment of GCN20 and a domain in GCN1 related to translation elongation factor 3 (EF-3) mediate complex formation between the two proteins. GCN20 contains ATP-binding cassettes found in membrane transporters; however, these sequences are partially dispensable for GCN20 function, and GCN1 and GCN20 are distributed uniformly throughout the cytoplasm. Based on their homology to EF-3, we propose that GCN1/GCN20 interact with ribosomes and facilitate binding of uncharged tRNA to GCN2. GCN2 contains a pseudo-kinase domain in its N-terminus that is conserved in the *Neurospora* homologue of GCN2 that is also required for regulating kinase function in vivo. The transcriptional activation domain of GCN4 contains 7-8 subdomains, each consisting of 2-3 bulky hydrophobic amino acids surrounded by acidic residues. These subdomains can cooperate in different combinations to activate transcription. An *ADE5,7* promoter fragment sufficient for adenine-repressible transcription contains three 6-10nt elements that probably function as binding sites for the BAS1, BAS2, and ABF1 transcription factors. Overexpressing BAS2 increases *ADE5,7* expression under repressing conditions, suggesting that adenine repression involves reducing the ability of *ADE* gene promoters to compete with other yeast genes for binding limiting amounts of BAS2.

Project Description:Objectives:

To understand at the molecular level how the yeast *S. cerevisiae* regulates its capacity for synthesizing amino acids and nucleotides according to the availability of these nutrients in the environment. One facet of this regulation, known as general amino acid control, involves increased expression of a large number of amino acid biosynthetic genes in response to starvation for any amino acid. Synthesis of GCN4 protein, the transcriptional activator in this system, is stimulated under starvation conditions by a translational control mechanism involving short open reading frames (uORFs) in the *GCN4* mRNA leader, several general translation initiation factors, and a protein kinase known as GCN2. Genetic and biochemical experiments are being conducted to understand how the uORFs regulate the flow of scanning ribosomes to the *GCN4* start codon according to the availability of amino acids. A combination of genetic, molecular and biochemical approaches are being used to identify and characterize the trans-acting factors, both positive and negative, that mediate the regulatory functions of the uORFs. Some of these factors are subunits of the general translation initiation factors eIF-2 or its guanine nucleotide exchange factor eIF-2B; others appear to be components of previously unidentified general initiation factors; still others, including the protein kinase GCN2 and its positive effectors GCN1 and GCN20, are dispensable regulatory factors dedicated to *GCN4* control. We wish to identify the biochemical functions of each factor and determine its position in the signal-transduction pathway that detects uncharged tRNA in amino acid-starved cells and modifies the translational machinery in a way that stimulates *GCN4* expression. This regulatory mechanism involves phosphorylation of the α subunit of eIF-2 by the protein kinase GCN2 and consequent reduction in the ability of eIF-2B to recycle eIF-2. This same mechanism operates in mammalian cells to inhibit total protein synthesis in response to various stress conditions, including amino acid starvation. By combining the powerful genetics and molecular biology of yeast with the biochemical analysis of translation initiation, we hope to provide a detailed molecular description of this highly conserved mechanism for regulating protein synthesis. The human eIF-2 α kinase DAI (double-stranded RNA-activated inhibitor of translation) is an important regulator of cell growth and differentiation and a critical component of the interferon response to viral infection in humans. We have shown that DAI can functionally substitute for GCN2 in yeast cells and we are using the *GCN4* genetic system to probe the mechanism of DAI activation by dsRNA.

In addition to studying the regulation of *GCN4* translation, we are conducting an in-depth mutational analysis of the transcriptional activation domain of the GCN4 protein, with the goal of obtaining a better understanding of the detailed structure of this domain and the identification of specific transcription factors with which GCN4 interacts in stimulating the expression of its target genes. We have also undertaken a study of the transcriptional control of adenine biosynthetic genes in response to exogenous purines. We are defining the minimal cis-acting sequences at the *ADE5* gene necessary for adenine repression of this gene and studying the regulatory proteins that bind to these sites, including BAS1, BAS2 and GCN4. We hope to learn how the expression and activity of these regulatory proteins is modulated by purine availability in the cell. Parallel studies of general amino acid control and the regulation of nucleotide biosynthesis in yeast should provide an integrated view of how eukaryotic cells control the levels of substrates (amino acids and nucleotides) and the machinery for protein and nucleic acid synthesis in the face of a changing nutritional environment.

Major Findings

I. Mechanism of translational control of *GCN4* expression.

A. Evidence that the efficiency of *GCN4* translation is inversely coupled to the concentration of eIF-2•GTP•Met-tRNA^{Met}_i ternary complexes in the cell.

Phosphorylation of the α subunit of translation initiation factor-2 (eIF-2 α) on residue Serine-51 is a prominent mechanism for regulating protein synthesis in mammalian cells. The phosphorylated form of eIF-2 inhibits translation initiation by impairing the conversion of eIF-2-GDP to eIF-2-GTP by the guanine-nucleotide exchange factor eIF-2B. In yeast, phosphorylation of eIF-2 by the protein kinase GCN2 stimulates translation of *GCN4* mRNA under conditions of amino acid starvation. According to our model, essentially all ribosomes that bind to the 5' end of *GCN4* mRNA translate uORF1, under both starvation and nonstarvation conditions, and ca. 50% remain attached to the mRNA and resume scanning downstream as 40S subunits. Under nonstarvation conditions, when the active form of eIF-2 is abundant, these 40S subunits rapidly rebind the eIF-2•GTP•Met-tRNA^{Met}_i ternary complex and regain the ability to recognize an AUG codon as a translational start site. Consequently, most reinitiate at uORFs 2, 3, or 4, and then dissociate from the mRNA, failing to reach the *GCN4* start codon. Under starvation conditions, eIF-2 α is phosphorylated by the protein kinase GCN2 and, by analogy with mammalian systems, this reduces the level of eIF-2•GTP•Met-tRNA^{Met}_i ternary complexes by inhibiting the recycling factor eIF-2B. Following translation of uORF1, many ribosomes now scan the entire distance between uORF1 and uORF4 without rebinding the ternary complex. Lacking the initiator tRNA^{Met}, they cannot recognize the AUG start codons at uORFs 2, 3, and 4 and continue scanning downstream. While traversing the leader segment between uORF4 and *GCN4*, most of these ribosomes re-bind the ternary complex and reinitiate translation at *GCN4*.

One of the key tenets of our model is that increased translation of *GCN4* is triggered by a reduction in the concentration of eIF-2•GTP•Met-tRNA^{Met}_i ternary complexes. In support of this idea, we showed that reducing the number of chromosomal genes encoding initiator tRNA^{Met} mimics the effect of eIF-2 phosphorylation in causing derepression of *GCN4* translation in the absence of amino acid starvation or GCN2 function. In addition, overexpression of all three subunits of eIF-2, which leads to a ~10-fold increase in the heterotrimeric eIF-2 complex in vivo, prevents derepression of *GCN4* in response to eIF-2 phosphorylation by GCN2. It also partially suppresses the slow-growth phenotype associated with hyperphosphorylation of eIF-2 by GCN2^c enzymes, and this suppression is augmented when tRNA^{Met}_i is co-overexpressed with the eIF-2 complex. These results provide strong support for the idea that phosphorylation of eIF-2 inhibits general translation and specifically stimulates *GCN4* expression by reducing the levels of the ternary complex in the cell.

B. Saturation mutagenesis of the regulatory domains in the GCD2, GCD7 and GCN3 subunits of eIF-2B.

Another important piece of evidence supporting our model is that mutations can be isolated in subunits of eIF-2B that suppress the phenotypes associated with eIF-2 α hyperphosphorylation by hyperactivated GCN2^c protein kinases. The suppressor mutations include point mutations and deletions in *GCN3* a nonessential subunit, and point mutations in *GCD7* and *GCD2* two of the four essential subunits of the yeast eIF-2B complex. These mutations could decrease the affinity of eIF-2B for phosphorylated eIF-2 and thus prevent sequestering of eIF-2B in an inactive state; alternatively, they could allow eIF-2B to catalyze nucleotide exchange on eIF-2. The C-terminal half of GCD2, GCD7 and GCN3, are very similar in amino acid sequence, suggesting that the interaction of each protein with eIF-2 involves a structural feature that is common to all three. To test this idea, we set out to identify all of the amino

acids in *GCD2*, *GCD7* and *GCN3* that are critically required for negative regulation of eIF-2B by phosphorylated eIF-2. Towards this end, we have isolated and characterized a large number of dominant or semi-dominant point mutations in *GCD2*, *GCD7*, and *GCN3* with the suppressor phenotype described above. For both *GCD2* and *GCD7*, we have obtained point mutations that render phosphorylated eIF-2 completely ineffective in down-regulating general protein synthesis and in stimulating *GCN4* translation (*Gcn⁻* phenotype). These mutations are actually more effective than a deletion of *GCN3* in suppressing the toxic effect of eIF-2 hyperphosphorylation by PKR. Thus, it appears that all three proteins make important contributions to the inhibitory effects of eIF-2(α P) on eIF-2B activity.

The results obtained thus far identify two clusters of *Gcn⁻* mutations which affect residues located in regions of homology between *GCD2*, *GCD7* and *GCN3*. One of the two clusters in each protein is located at the extreme C-terminus, the domain of greatest sequence similarity among the three. A second cluster occurs within a region of high similarity near the N-termini of *GCN3* and *GCD7* and in the center of *GCD2* (which contains a large N-terminal domain not shared with *GCD7* and *GCN3*). It is noteworthy that all of the *GCD2* mutations fall within the region of the protein that is related in sequence to *GCD7* and *GCN3*. In a few cases, amino acids occupying identical positions in the sequence alignments have been mutated in two of the three proteins. Together, these results suggest that homologous segments in *GCD2*, *GCD7* and *GCN3* are devoted to the regulatory interactions between eIF-2B and phosphorylated eIF-2. These homologous segments might be juxtaposed on the surface of eIF-2B and make direct contact with different residues in the phosphorylated N-terminal domain of eIF-2 α , as suggested previously. Alternatively, one of the two regulatory segments in each protein might interact with eIF-2 α while the other could mediate a conformational change in eIF-2B that distorts the active site when phosphorylated eIF-2 is bound. The segments located between the two clusters of *Gcn⁻* mutations in each protein are also very similar in sequence. Perhaps these segments perform a structural role in mediating subunit interactions between *GCD2*, *GCD7*, and *GCN3*.

C. Suppression of eIF-2(α P) toxicity by overexpression of eIF-2B subcomplexes.

A second approach we have taken in probing the structure and function of eIF-2B is to overexpress different combinations of the eIF-2B subunits. We showed previously that overexpressing all 5 subunits, or all 4 subunits except *GCN3*, suppresses the toxicity of eIF-2 α phosphorylation in *GCN2^c* strains. More recently, we observed the same phenotype when only *GCD2*, *GCD7* and *GCN3* were overexpressed from high copy-number plasmids. In fact, overexpressing just *GCD7* and *GCD2* conferred nearly the same suppression as did the combination of all three proteins. In contrast, overexpressing each of the individual subunits and many other combinations of two or three subunits of eIF-2B conferred no suppression of *GCN2^c* mutations. One interpretation of these results is that *GCD7*, *GCD2* and *GCN3*, or just *GCD7* and *GCD2* alone, can form stable subcomplexes that sequester phosphorylated eIF-2 and neutralize its inhibitory effect on the native eIF-2B complex. In support of this interpretation, we succeeded in co-immunoprecipitating nearly all of the excess *GCD7* and much of the excess *GCN3* with anti-*GCD2* antibodies from cells overexpressing *GCD2*, *GCD7* and *GCN3*. Thus, we now have biochemical evidence for subcomplex formation by these proteins *in vivo*. An alternative explanation for our results is that subcomplexes containing *GCD2*, *GCD7* and *GCN3* have nucleotide exchange activity and thus can functionally replace native eIF-2B. This explanation seems unlikely because overexpressing the subcomplex does not complement the growth defects of mutations in the *GCD1* subunit that are suppressed when eIF-2 is overexpressed.

D. In vivo analysis of the mechanism for inhibiting eIF-2B by phosphorylated eIF-2.

Overexpression of eIF-2 in strains containing an activated GCN2^c kinase leads to the production of much higher levels of the phosphorylated form of the protein than are present in the same strain containing wild-type amounts of eIF-2; yet, we found that the toxicity of GCN2^c proteins is suppressed by overexpressing eIF-2. This apparent paradox can be explained by noting that the ratio of phosphorylated to nonphosphorylated eIF-2 was reduced in the strain overexpressing eIF-2 versus the parental strain. If we propose that the mechanism of inhibition of eIF-2B by eIF-2(α P) is competitive inhibition, rather than irreversible binding of a non-competitive inhibitor, we can explain why the degree of inhibition depends more on the ratio of eIF-2(α P):eIF-2 rather than the absolute amount of phosphorylated eIF-2. We can eliminate the trivial possibility that overexpression of eIF-2 simply decreases the requirement for eIF-2B, perhaps as a result of spontaneous nucleotide exchange on eIF-2, because several mutations in subunits of eIF-2B are not complemented by overexpression of eIF-2. Therefore, our results indicate that eIF-2(α P) acts as a competitive inhibitor of eIF-2B rather than forming an extremely stable inactive complex with it.

E. GCD10 is the RNA binding subunit of eIF-3.

Mammalian eIF-3 is an 8-subunit complex that stimulates several steps in the initiation pathway in cell-free translation systems. Yeast contains a structurally similar complex that can functionally replace mammalian eIF-3 in an in vitro translation system containing all the initiation factors from HeLa cells except for eIF-3. In collaboration with Mercedes Tamame's and John Hershey's laboratories, we have now obtained strong biochemical evidence that *GCD10* encodes the 54.6kDa RNA-binding subunit of yeast eIF-3. Mutations in *GCD10* lead to constitutive derepression of *GCN4* translation in the absence of *GCN2* and *GCN3*, and to temperature-sensitive growth on rich medium, the same phenotypes associated with *gcd* mutations affecting subunits of eIF-2 or eIF-2B. We cloned *GCD10* and found that a deletion of the gene is lethal. In addition, we observed dissociation of polysomes and accumulation of 80S subunits after shifting *gcd10* temperature-sensitive mutants to the non-permissive temperature, consistent with a defect in an essential translation initiation factor. After producing antibodies against GCD10, we showed that the protein is present in a high molecular weight complex, and that a fraction of GCD10 is physically associated with polysomes and ribosomal subunits. GCD10 co-purifies with eIF-3 biochemical activity and other subunits of the eIF-3 complex through Superose-6 gel filtration chromatography and MonoS ion-exchange chromatography, and is identical in electrophoretic mobility to the 62 kDa subunit of eIF-3 described previously. This 62 kDa subunit binds RNA in vitro and we showed that GCD10 has RNA binding activity by a Northwestern assay using radiolabeled globin mRNA as the probe. We also discovered that our GCD10 antibodies cross-react with the RNA-binding 66kDa subunit of human eIF-3. Finally, we showed that GCD10 is specifically co-immunoprecipitated with the ~90 kDa subunit of eIF-3 encoded by *PRT1*.

Biochemical studies on mammalian eIF-3 have implicated this factor in several steps of initiation, including dissociation of 80S ribosomes into free 60S subunits and 46S preinitiation complexes containing eIF-3 and 40S subunits, binding of eIF-2/GTP/Met-tRNA^{Met}_i ternary complexes to form 43S complexes, and binding of mRNA to 43S complexes containing both eIF-3 and the ternary complex. We have proposed that *gcd10* mutations reduce the ability of eIF-3 to rebind to 40S subunits and form 46S complexes following termination at uORF1. This would result in "naked" 40S subunits scanning downstream from uORF1 which have a reduced ability to rebind ternary complexes (Fig. 8). Alternatively, *gcd10* mutations could decrease the ability of eIF-3 to stimulate binding of eIF-2/GTP/Met-tRNA^{Met}_i ternary complexes to 46S complexes rather than delaying the formation of these complexes. Either of these last two defects would decrease the rate at which ternary complexes can

rebind to 40S subunits scanning downstream from uORF1, explaining why *gcd10* mutations cause ribosomes to ignore uORFs 2-4 and reinitiate at *GCN4* in the absence of eIF-2 phosphorylation by GCN2.

It is thought that eIF-3 functions at multiple steps of the initiation pathway and interacts with most of the other known initiation factors, including eIF-1A, eIF-2A, eIF-3A, eIF-4, eIF-4A, eIF-4B, and eIF-5. It could thus be imagined that eIF-3 provides a platform on the 40S subunit which facilitates binding of these factors in the correct orientation with respect to one another and with the decoding sites on the ribosome. One way of testing this idea is to employ genetic suppressor analysis to identify specific interactions between particular subunits of eIF-3 and the components of other initiation factors or ribosomal proteins. Towards this end, we have isolated high copy plasmids from wild-type yeast genomic libraries that complement temperature-sensitive *gcd10* mutations. Five of the six plasmids partially suppress the derepression of *GCN4* (Gcd⁻ phenotype) seen in *gcd1*, *gcd2* and *gcd13* mutants in addition to *gcd10* mutants; however, none suppresses the temperature-sensitive phenotypes of these other *gcd* mutations, or of *prt1* mutations, suggesting a specific interaction with GCD10. Hybridization of the cloned fragments to the Olson-Riles ordered genomic library shows that the suppressor genes derive from seven genomic loci. Five of these loci contain one of the structural genes for initiator tRNA^{Met} and in one case, we have shown that disruption of this gene abolishes suppressor activity. Thus, it appears that overexpression of initiator tRNA^{Met} can compensate for the general initiation defect in *gcd10* mutants. The simplest explanation for this finding is that binding of the ternary complex to small ribosomal subunits is defective in the *gcd10* mutants. This provides in vivo evidence that eIF-3 stimulates binding of ternary complexes to 40S subunits and supports the idea that *gcd10* mutations are specifically defective for this aspect of eIF-3 function. The fact that multi-copy tRNA^{Met} genes do not suppress mutations in *PRT1* encoding another subunit of eIF-3, seems to indicate that *PRT1* mutations alter this or some other function of eIF-3 in a way that cannot be overcome by simply increasing the concentration of ternary complexes.

F. Analysis of novel GCD factors

Mutations in the *GCD13* and *GCD14* genes have the same phenotypes as mutations in any of the four GCD genes encoding subunits of eIF-2B. Thus, GCD13 and GCD14 could be additional subunits of eIF-3 or components of another factor that influences the production or utilization of the ternary complex in translation initiation. In accordance with this idea, we found that *GCD13* mutations lead to polysome dissociation in vivo when mutants are incubated at the nonpermissive temperature. Efforts to clone *GCD13* by conventional approaches have failed; therefore, we mapped *GCD13* to the telomeric region of the left arm of chromosome XV. Using this information we are attempting to isolate the gene from ordered lambda clones or cosmids containing this portion of XV.

We confirmed that *GCD14* mutations lead to constitutive derepression of *GCN4* expression in the absence of GCN2 and GCN3. The *GCD14* gene has been isolated and found to encode a protein of ca. 40 kDa that contains a degenerate RNA recognition motif (RRM), and thus could be an RNA binding protein. Based on peptide sequencing done in John Hershey's lab, GCD14 does not appear to be the 39kDa subunit of eIF-3.

G. GCN1 and GCN20 are components of a heteromeric protein complex that stimulates the phosphorylation of eIF-2 α by GCN2.

GCN1 and GCN20 are positive regulators of *GCN4* translation that function by stimulating the ability of GCN2 to phosphorylate eIF-2 in vivo under amino acid starvation conditions. We have shown that GCN1 and GCN20 interact in vitro by co-immunoprecipitation experiments and by using the yeast two-

hybrid system; thus, they are components of a protein complex that is required *in vivo* to couple GCN2 kinase activity to amino acid availability. Interestingly, ca. 800 amino acids of GCN1 shows sequence similarity to the fungal-specific translation elongation factor EF-3. It is thought that EF-3 functions in stimulating release of uncharged tRNA from the E (exit) site of the ribosome and thereby stimulates binding of charged tRNA to the A site. Remarkably, GCN20 also shows strong sequence similarity to a portion of EF-3 containing two ATP-binding cassettes characteristic of the "ABC" family of transporter proteins. GCN1, in contrast, does not contain the signature sequences of ABC domains and is most similar to EF-3 in a region N-terminal to its ABC domains. Thus, GCN20 and GCN1 are similar to different portions of EF-3.

The great majority of ABC proteins are membrane transporters that use the energy of ATP hydrolysis to pump substrates against a concentration gradient. The typical ABC transporter, typified by the multiple-drug resistance protein MDR, consists of four membrane-associated domains. Two of these domains contain six hydrophobic α -helical membrane-spanning segments that form the pore through which substrates cross the membrane. The other two domains are peripherally located on the cytoplasmic face of the membrane, bind ATP and couple ATP hydrolysis to the transport process. GCN20 seems to lack regions sufficiently hydrophobic to function as transmembrane domains; however, in prokaryotes at least, the individual domains of the transporter can be expressed as separate polypeptides. It is thus conceivable that GCN20 contains the two ATP-binding domains and interacts with one or more transmembrane proteins in carrying out a transport function. One possibility would be that GCN20 interacts with GCN1, which does contain numerous hydrophobic regions, to comprise one of the transporters identified in yeast that delivers amino acids from the cytoplasm to the vacuole.

A few members of the large family of ABC proteins, such as EF-3, are not associated with membrane transport events. Therefore, GCN20 and GCN1 could be components of an ABC complex with no role in membrane transport that acts more directly to stimulate GCN2 kinase activity by uncharged tRNA. One possibility, motivated by their sequence similarity to segments of EF-3, would be that GCN1 and GCN20 interact with the ribosome and facilitate an interaction between uncharged tRNA bound at the ribosomal A site and the HisRS-related domain of GCN2. This activity could be akin to the proposed function of EF-3 in stimulating release of uncharged tRNA from the ribosomal E site. We have identified proteins from diverse eukaryotic organisms, including *C. elegans*, rice, Arabidopsis, and humans, that are more closely related to GCN20 than to any other known ABC proteins, and contain sequence similarity in regions flanking the ABC domains that are not conserved among diverse members of the ABC family. These sequence similarities could indicate that GCN20 belongs to a new subfamily of ABC proteins with a biological function conserved from yeast to mammals. If so, higher eukaryotes may contain a mechanism for coupling the rate of translation initiation to the level of uncharged tRNA similar to that uncovered in *S. cerevisiae* involving GCN2 and its positive effectors GCN1 and GCN20.

We have used our antibodies against GCN1 to localize the protein in yeast cells by indirect immunofluorescence. Our results to date indicate that the protein is uniformly distributed throughout the cytoplasm. Its distribution appears distinctly different from that of a known vacuolar membrane protein, VAT2, that was analyzed in parallel. These results do not support the model that the GCN1/GCN20 complex are components of a membrane transporter. We have also been investigating the importance of the ABC domains in GCN20 for its regulatory function by making deletions and site-directed substitutions of conserved residues in the Walker A and B motifs of the ATP-binding cassettes. Surprisingly, mutations in these motifs reduce but don't eliminate GCN20 function. In fact, a deletion allele lacking all 634 residues C-terminal to amino acid 118, which includes both

ABC domains, retains substantial complementing function. In contrast, a deletion of residues 4-118 that leaves both ABC domains intact completely abolishes GCN20 function. These mutations do not significantly lower GCN20 protein levels. The N-terminal 118 residues is sufficient for interaction with the C-terminal two-thirds of GCN1 in the yeast two-hybrid system. In addition, a GCN20-lacZ fusion protein containing only the N-terminal 118 residues of GCN20 can be co-immunoprecipitated with GCN1, whereas a nearly full-length GCN20-lacZ fusion lacking only residues 4-118 fails to interact with GCN1. Thus, the N-terminal domain of GCN20 is both necessary and sufficient for its function as a positive regulator of GCN2 and for its physical interaction with GCN1. The fact that the ABC domains of GCN20 are partly dispensable for its positive regulatory function also seems to be at odds with the idea that GCN1 and GCN20 form an amino acid transporter because ATP hydrolysis is essential for the function of ABC transporter proteins. However, we found a predicted ABC protein highly similar to GCN20 (Yer036), in the yeast genome database that could conceivably substitute for the ABC domains of GCN20 in a *gcn20* mutant that retains only the N-terminal 118 residues.

Using the yeast two-hybrid system, we showed that the N-terminal 118 amino acids of GCN20 could interact with the C-terminal two-thirds of GCN1. We have been extending this analysis by testing different overlapping segments of GCN1 for interaction with the N-terminus of GCN20. Our results thus far indicate that all GCN1 fragments that interact with GCN20 contain the EF-3-related segment of GCN1. The smallest of these segments contains 345 amino acids and is coincident with the region most highly related to EF-3.

H. GCN2-independent derepression of *GCN4* translation.

The regulatory mechanism described above applies to the derepression of *GCN4* translation that occurs under conditions of prolonged amino acid-limited growth. There is evidence that a different pathway is responsible for the transient derepression of *GCN4* that accompanies a "shiftdown" from amino acid-rich to minimal medium. We uncovered another instance of *GCN2*-independent induction of *GCN4* translation in strains overexpressing tRNAs under conditions in which the excess tRNA cannot be aminoacylated. The most complete derepression occurred in strains overexpressing a mutant tRNA^{Val}(AAC) with a mutation in the acceptor stem that is expected to abolish aminoacylation. Overexpression of the mutant tRNA^{Val} derepresses *GCN4* translation under nonstarvation conditions in a manner dependent on the uORFs and independent of *GCN2* and Ser-51 on eIF-2 α . Overexpression of the mutant tRNA^{Val} did not affect cellular growth under nonstarvation conditions in an otherwise wild-type strain; however, it exacerbated the slow-growth phenotype of a *GCN2*^c mutant. These findings suggest that the presence of excess uncharged tRNA^{Val} lowers the level of eIF-2 activity by a mechanism that does not involve eIF-2 α phosphorylation. A weaker *GCN2*-independent derepression was observed in *gcn2* Δ strains overexpressing wild-type tRNA^{His} under histidine starvation conditions, indicating that the *GCN2*-independent mechanism can be activated by an excess of uncharged wild-type tRNA.

II. Regulation of the mammalian eIF-2 α kinase DAI in yeast.

In mammalian cells, stimulation of the latent protein kinase activity of the double-stranded-RNA-activated inhibitor of translation, DAI, occurs in response to viral infection. The N-terminus of DAI contains two copies of a sequence motif found in several dsRNA-binding proteins and numerous deletion and point mutations have been introduced into these sequences that affect dsRNA binding in vitro. We showed that DAI expressed at low levels in yeast mimics the function of GCN2 in stimulating translation of *GCN4*, and inhibits total protein synthesis when expressed at high levels. Exploiting these findings, we went on to demonstrate that the dsRNA-binding domain is required for

DAI kinase function in yeast cells and that the dsRNA-binding motifs are involved in the activation mechanism. Secondly, we find that the N-terminal copy of the dsRNA-binding domain plays a greater role than the C-terminal copy in activating kinase function in yeast cells. Thus, the requirements for dsRNA binding in vitro and for kinase activation in vivo closely coincide, providing direct evidence that dsRNA binding to the repeated motifs stimulates DAI kinase activity in vivo.

We have been attempting to identify sites of autophosphorylation in PKR that are important for its activation by dsRNA in vivo. Mike Mathew's lab have identified several tryptic peptides from the N-terminal regulatory domain of PKR that are autophosphorylated in vitro by PKR purified from cultured human cells. In collaboration with their group, we have analyzed the effects of making alanine substitutions in the various Ser and Thr residues in these peptides on PKR function in yeast. Our results thus far indicate that substitutions at a combination of these potential phosphorylation sites partially impairs PKR activation. This result implies that, although these sites may be important for activation, there are probably additional sites of autophosphorylation in the catalytic domain of PKR. Accordingly, we have also been examining the importance of two Thr residues located in the insert between kinase subdomains VII and VIII as potential autophosphorylation sites. The activity of several protein kinases, including PKA, MAPK, MAPKK and CDK requires phosphorylation or autophosphorylation of one or two closely spaced Ser or Thr residues in this segment. In the case of CDK, this phosphorylation is thought to alter the conformation of a loop that protrudes into the active site and prevents proper positioning of the protein substrate. It can also promote proper interactions between the N-terminal and C-terminal lobes of the kinase domain, as in the case of PKA and MAPK. We have preliminary evidence that Ala substitutions at Thr-445 and Thr-450 each decrease PKR kinase function in vivo, and completely inactivate it when combined in the same protein. These mutant proteins are expressed at high levels and migrate in SDS-PAGE with the same mobility as the subdomain II K296R mutant protein; thus, the mutations at Thr-445 and Thr-450 appear to prevent autophosphorylation of PKR on most, if not all of its phosphorylation sites.

III. Mechanism of transcriptional activation by GCN4.

A. Mutational analysis of the GCN4 activation domain.

Previous work by Struhl and his colleagues established that GCN4 contains a potent transcriptional activation domain located in the center of the protein between residues 107-144. There were indications that additional activation determinants were located N-terminal to this central acidic activation domain (CAAD); however, they appeared to be insufficient for high level activation in the absence of a portion of the CAAD. We set out to locate more precisely the activation determinants in the N-terminal portion of GCN4 and to compare the relative efficiencies of transcriptional activation by the CAAD versus the N-terminal activation domain (NTAD) when GCN4 is expressed at physiological levels. Towards this end, numerous deletion and point mutations were constructed in a single-copy *GCN4* gene containing the native promoter and translational control elements, and analyzed for their effects on gene expression after inducing the mutant proteins by histidine starvation. The levels of all mutant proteins were quantified by immunoblotting using antibodies raised against the DNA-binding and dimerization domain of GCN4. Because activation domains in several well characterized mammalian activators, most notably VP16 and yeast GAL4, contain hydrophobic residues critically required for the activation functions of these proteins, we also wished to determine whether this structural feature extended to the CAAD or NTAD of GCN4 protein.

Our analysis showed that GCN4 contains a potent activation domain in its N-terminal 100 amino acids that confers nearly wild-type transcriptional activation of *HIS3* or *HIS4* in the complete absence of the CAAD. Similarly, the CAAD can promote high-level transcription when the NTAD is missing. The

CAAD and NTAD are both strongly dependent on the ADA2 protein for activating transcription at *HIS3* and *HIS4*. The CAAD was found to be critically dependent on two clusters of aromatic and bulky hydrophobic residues: Met-107, Tyr-110, Leu-113 and Trp-120, Leu-123 and Phe-124. The latter corresponds to a short stretch of amino acids (Trp-X-Ser-Leu-Phe), that is conserved between the activation domains of GCN4 and its homologues in *Neurospora* (*cpc-1*) and *Aspergillus* (G. Braus, personal communication). The NTAD was found to be dependent on a pair of Phe residues at positions 97-98. Consistent with the deletion analysis, point mutations in the hydrophobic and aromatic residues in the CAAD impair GCN4 function only when combined with a deletion of the NTAD or with substitutions at Phe-97, Phe-98. Likewise, deletions in the NTAD or point mutations at Phe-97 and Phe-98 impair GCN4 function only when the CAAD is missing or bears mutations at the aforementioned hydrophobic and aromatic residues located between Met-107 and Phe-124. Combining together substitutions at all eight hydrophobic residues in full-length GCN4 essentially eliminated activation of *HIS3* and reduced activation of *HIS4* to 20% of wild-type. These findings indicate that the activation domain of GCN4 is more complex than was previously imagined and shares significant structural similarities with activation domains found in other acidic activators.

In accordance with the results just described, a deletion that removes the entire CAAD and Phe-97 and Phe-98 does not completely abolish activation of *HIS4*. Surprisingly, overexpression of this mutant protein achieved by removing the uORFs is lethal under nonstarvation conditions. This lethality is dependent on the GCN4 DNA binding domain, suggesting that it results from sequestration of one or more general transcription factors in non-productive complexes by the excess mutant GCN4 protein.

Based on these observations, we set out to identify activation determinants in the NTAD located in the region N-terminal to position 97. To accomplish this goal, PCR was employed in a random mutagenesis of the NTAD between residues 17-100 in a *GCN4* allele lacking the CAAD, and alleles with greatly reduced GCN4 function were isolated and sequenced. Thirty out of 42 alleles analyzed contained a mutation at one of 6 Phe residues present in the NTAD segment that we mutagenized (Phe-45, Phe-48, Phe-67, Phe-69, Phe-97, and Phe-98). Using site-directed mutagenesis, we found that single alanine substitutions of these Phe residues could account for the *Gcn⁻* phenotype of all 30 alleles.

In one of the 12 mutant alleles in which a Phe residue was not altered, the mutant phenotype resulted from a lysine substitution at Glu-88. This was the only mutation at an acidic residue that had a strong effect on GCN4 function. Moreover, we found that an alanine substitution at Glu-88 has little effect on activation, indicating that an acidic residue is not required even at this position. Interestingly, Glu-88 maps within the only short stretch of amino acids in the NTAD that is conserved between GCN4 and its relatives in *Neurospora* and *Aspergillus*: Thr⁸²-X-Leu-X-X-Pro-X-Leu⁸⁹, located 8 residues N-terminal to the critical Phe-97, Phe-98 pair. Using site-directed mutagenesis, we found that Leu-89 is the sole conserved residue in this segment that is required for activation by GCN4. Thus, the Lys substitution at Glu-88 probably reduces activation by interfering with the function of the adjacent residue, Leu-89. It is noteworthy that leucine has a bulky hydrophobic side-chain and can partially substitute for certain critical Phe residues in the VP16 activation domains.

Two additional Phe residues are present in the N-terminal domain of GCN4 (Phe-9 and Phe-16) just outside the region that we mutagenized by PCR. Using site-directed mutagenesis, we found that Phe-9 is also critical, whereas Phe-16 makes only a minor contribution to NTAD function. A methionine at position 14 is also dispensable for activation by the NTAD. The latter results show that not all bulky hydrophobic or aromatic residues in the NTAD are functionally important. This conclusion was underscored by the fact that alanine substitutions at Leu-65 and Leu-71 (located very close to the critical Phe-67 Phe-69 pair) and at Leu-84, Val-93 and Val-94 (in the vicinity of Phe-97 Phe-98), have

no effect on NTAD function. To determine whether Phe is uniquely required at residue 45, we randomized this codon and isolated alleles covering a wide range of GCN4 function. Sequencing these alleles indicated the following functional hierarchy for different amino acids at position 45:

Trp, Phe > Leu > Tyr > Ile > Val, Ser > Thr, Pro, Glu, Gln, Cys, Ala, Lys, Gly, with the first group conferring wild-type activation and the last showing the null phenotype. Thus, either an aromatic or bulky hydrophobic residue seems to be required at position 45. Finally, we verified that alanine substitutions at the three new Phe clusters in the NTAD at positions 9/16, 45/48 and 67/ 69 reduce activation by full-length GCN4 only when combined with point mutations in the CAAD at Met-107, Tyr-110, Leu-113 and Trp-120, Leu-123 and Phe-124. In addition, we used immunoblot analysis to verify that these mutations do not reduce the level of GCN4 protein in the cell.

Taken together, our results indicate that GCN4 contains a minimum of 7 clusters containing 2-3 aromatic or bulky hydrophobic amino acids distributed throughout the N-terminal 150 amino acids that comprise its two activation domains. We cannot determine from our data whether these residues are required for the overall structure of the activation domains or for specific contacts they make with co-activators or basic transcription factors. It is worth noting, however, that certain transcription factors contain critical hydrophobic residues, including TBP and TAF_{II}230, and it has been suggested that hydrophobic interactions are important in stabilizing contacts between activators and their targets in the transcriptional machinery.

IV. Genetic analysis of the regulation purine biosynthetic genes.

A. Identification of the minimal cis-acting elements in the *ADE5,7* promoter required for adenine-repressible transcription.

The transcription of genes encoding enzymes of de novo purine nucleotide biosynthesis, is repressed by addition of adenine to the medium. The BAS1 and BAS2 proteins are required for high-level expression of *ADE5,7*, *ADE8*, *ADE1* and *ADE2* under derepressing conditions (minimal medium lacking adenine), and it is believed that adenine repression involves down-regulating the ability of BAS1 or BAS2 to stimulate transcription. One way this could occur would be if a repressor protein binds to the *ADE* genes when excess adenine is present in the medium. According to this model, a negative regulatory site would be present in the promoters and its removal would lead to constitutively derepressed expression. To test this possibility, we set out to identify the minimal sequence element from *ADE5,7* sufficient for adenine-regulated transcription. We found that a 74 bp fragment containing two TGACTC elements separated by 33 bp confers BAS1/2-dependent, adenine-repressible transcription on a *CYC1-lacZ* construct that contains a TATA box but lacks the native *CYC1* UAS elements. The gene-proximal TGACTC sequence in this fragment is required for promoter function and purine regulation, whereas the distal element augments expression that is fully dependent upon the proximal site. Interestingly, the proximal TGACTC sequence is more critical than the distal sequence at *ADE2* as well. Unexpectedly, deletion of the region between the two TGACTC sequences decreased expression dramatically from the *ADE5,7-CYC1* hybrid construct.

To map in greater detail the sequence requirements for this regulatory element, we are making clustered substitutions in consecutive 3-bp intervals extending across a 67 bp fragment that contains only the proximal TGACTC sequence and ca. 30 5'- and 3'-flanking nucleotides. From the results obtained thus far, we find that substitutions in the TGACTC core element and in the 6 nt immediately 3' to this sequence (GTGTCC) eliminate both expression and adenine regulation of the hybrid promoter, presumably due to inactivation of the BAS1 binding site. Changes in the sequence TAATAA located 10 nt 3' to the TGACTC site also greatly reduce expression and adenine regulation. These latter

nucleotides might also be required for BAS1 binding *in vivo*; alternatively, they may comprise a BAS2 binding site. Consistent with a requirement for sequences located between the two TGACTC elements, substitutions in the sequence CCGTCGGTAGTGACA located 24 nt 5' of the proximal TGACTC sequence either reduce or eliminate both expression and adenine regulation. Inspection of this sequence and consideration of the relative effects of mutations at different positions in the sequence strongly suggests that it constitutes a binding site for ABF1 protein. None of the ca. 50 clustered substitution mutants we analyzed thus far have the derepressed phenotype expected from inactivation of a negative control site where a repressor protein would bind. Therefore, our results are most consistent with the idea that adenine repression of transcription involves antagonizing the function of the positive factors BAS1, BAS2 or ABF1, through a protein modification like phosphorylation or by a protein-protein interaction with a repressor protein. BAS1 is the most likely target for regulation because both ABF1 and BAS2 interact with promoters that are presumably unresponsive to purine levels.

B. Genetic selections for trans-acting mutations that abolish adenine-mediated repression of transcription.

We wish to identify additional factors that might be involved in regulating BAS1 or BAS2 by adenine and to delineate the regulatory domains in BAS1 or BAS2. Towards these ends, we have constructed yeast strains in which defects in adenine repression lead to an easily detectable growth phenotype. These strains contain a fusion of the *ADE5,7* promoter to the *HIS3* coding region replacing the wild-type *HIS3* gene. These strains are His⁺ on minimal medium but His⁻ when grown on of adenine, due to repression of the *ADE5,7-HIS3* construct. By selecting for His⁺ revertants on adenine-containing medium, we hope to isolate mutations that abolish adenine repression of *ADE5,7*. These strains also contain an integrated *ADE5,7-lacZ* construct to identify mutations that derepress *ADE5,7* in *trans*, an *ade2* allele to eliminate adenine uptake mutants that are derepressed because of low intracellular adenine pools, and a *gcn4Δ* allele to avoid isolation of *gcd* mutations. We have begun using these strains to select three kinds of regulatory mutations.

First, we mutagenized a plasmid-borne copy of *BAS1* by hydroxylamine and selected two dominant alleles that allow growth of the test strain on +Ade -His medium. The products of these *BAS1* alleles may be insensitive to adenine repression or may have increased affinity for DNA or other proteins with which BAS1 interacts to activate transcription. It is unlikely that they simply increase BAS1 protein levels because wild-type *BAS1* on a high copy plasmid does not confer a His⁺ phenotype in the test strain. These two alleles have a relatively weak phenotype; therefore, we are screening a second pool of *BAS1* plasmids mutagenized with an *E. coli* mutator strain to obtain alleles with a stronger His⁺ phenotype. In addition, we will sequence the two existing mutations and construct a double mutation. We will also screen a pool of *BAS2* plasmids mutagenized with the mutator strain for the same phenotype.

In a second approach, we screened high copy plasmid yeast genomic libraries for dosage suppressors. Among the His⁺ transformants analyzed thus far, the only suppressor plasmid we obtained (that does not contain *HIS3*) encodes *BAS2*. As just mentioned, high copy-number *BAS1* does not suppress the His⁻ phenotype of the tester strain (although we have not confirmed that BAS1 is overexpressed from this plasmid). These findings could indicate that BAS2 is rate-limiting for activation of *ADE5,7* transcription, in accordance with the fact that BAS2 is involved in regulating other genes besides *ADE* genes. It also suggests a model in which BAS1 must recruit BAS2 to an *ADE* promoter to activate transcription, and this function of BAS1 is diminished by excess adenine. We are now determining whether overexpression of BAS2 leads to constitutively derepressed *ADE5,7-lacZ* expression, as

predicted by the model just described, or whether *ADE5,7* expression is elevated under derepressing conditions but remains adenine-repressible even when *BAS2* is overexpressed. The latter would indicate that recruitment of *BAS2* from other promoters is not the adenine-repressible parameter.

In a third approach, we have isolated chromosomal mutations that suppress the His⁻ phenotype of the test strain on +Ade -His medium. Thus far, we obtained 10 His⁺ revertants in which expression of the *ADE5,7-lacZ* fusion is derepressed 2-3 fold on medium containing adenine, indicating the presence of a trans-acting mutation that impairs adenine repression. We are investigating whether this phenotype is conferred by a single mutation in each revertant and determining the number of complementation groups involved. Once this analysis is completed, we will clone the wild-type alleles of the suppressor genes by complementing the His⁺ phenotype of the mutants. It will also be worthwhile to carry out insertional mutagenesis with the Snyder library using the tester strains described above and selecting for His⁺ mutants on +Ade -His medium. If the key regulatory factors are nonessential, this approach should allow rapid identification and isolation of the genes involved.

Significance to Biomedical Research and the Program of the Institute:

Yeast is a eukaryotic organism that carries out many of the fundamental processes of eukaryotic cells and offers a level of molecular genetic analysis that cannot be achieved with multicellular organisms. General amino acid control and the regulation of purine biosynthetic genes involve a variety of strategies for regulating gene expression. Investigating the molecular details of these two systems is providing useful paradigms for coordinate transcriptional activation of unlinked genes, transcriptional repression, gene-specific translational control, functional modifications of protein complexes involved in general translation initiation, and the involvement of protein kinases and protein phosphatases in signal-transduction pathways. The results of these inquiries can be expected to provide valuable insights into how higher eukaryotic cells control the availability of amino acid and nucleotide substrates for protein and nucleic acid synthesis in response to changes in the environment. In fact, *GCN4* translational control appears to be a gene-specific instance of a highly conserved mechanism that regulates protein synthesis in mammalian cells under conditions of stress such as amino acid starvation. The regulatory mechanisms under study in yeast are critical elements of the network of interrelated controls that couple growth rate and division to environmental stimuli in all eukaryotic cells.

Proposed Course of the Project.

A. Investigation of the role of poly-A binding protein (PAB) in reinitiation on *GCN4* mRNA.

Previous studies have suggested that inactivation of PAB impairs general translation initiation and that this defect can be partially overcome by a variety of mutations that reduce 60S subunit biogenesis. This has led to the idea that PAB is involved in some aspect of 40S-60S subunit joining; however, its precise role in translation initiation is unknown. We propose to isolate mutations in PAB that perturb *GCN4* translational control as a means of increasing our understanding of PAB function in translation. There are three principal ways that perturbing PAB function could affect *GCN4* expression. First, it could decrease the ability of ribosomes to remain attached to the mRNA, or to resume scanning and re-form an initiation complex, following termination at uORF1. This would produce a Gcn⁻ phenotype. Second, it could decrease the dissociation of ribosomes from the mRNA following termination at uORF4, causing a Gcd⁻ phenotype. A third possibility is that PAB could be required for efficient subunit joining during reinitiation. If PAB must rebind to ribosomes as they scan downstream from uORF1, as postulated for the eIF-2·GTP·Met-tRNA^{Met} ternary complex, then reduced-function mutations in PAB should have a greater effect on reinitiation at uORF4 than at *GCN4* and produce a Gcd⁻ phenotype. They would lead to increased *GCN4* expression only when uORF1 and uORF4 are

both present in the leader, whereas a Gcd⁻ phenotype arising from reduced ribosome dissociation following uORF4 translation (the second mechanism above) would be seen with uORF4 alone.

To isolate mutations in PAB affecting *GCN4* translation, we will mutagenize the cloned gene on a plasmid, subclone various segments into a wild-type copy of *PAB* and analyze the resulting plasmids in a *pab1Δ* strain as the only copy of *PAB* by plasmid shuffling. Mutant plasmids will be screened for altered sensitivity to 3-AT or 5-FT to identify Gcn⁻ or Gcd⁻ phenotypes, respectively. It will be particularly interesting if mutations having different effects on *GCN4* are found to map in different domains of PAB. This mutational analysis should also be helpful in identifying portions of the protein that are specifically involved in translation initiation as opposed to other potential functions of PAB in poly(A) tail metabolism and mRNA turnover. In addition, mutations with a 3-AT^s phenotype could be used in suppressor analysis to identify proteins that interact with PAB in the translation initiation pathway.

B. Molecular genetic analysis of eIF-2B.

1. In vitro analysis of protein-protein interactions in the GCD2/GCD7/GCN3 subcomplex.

From immunoprecipitation experiments, we obtained biochemical evidence that GCD2, GCD7 and GCN3 can form a stable subcomplex in vivo. To test our prediction that the GCD2/GCD7/GCN3 subcomplex can interact with eIF-2, we intend to purify the subcomplex from yeast, or reconstitute it in vitro from proteins overexpressed in *E. coli*, and probe for interactions with eIF-2 purified from yeast. With the higher protein concentrations that can be achieved with purified preparations, we hope to detect complexes that are not stable enough to be detected by immunoprecipitations from cell extracts. We have constructed plasmids for overexpression of GCD2, GCD7 and GCN3 in *E. coli* and established conditions for purification of these proteins from bacteria. We can now purify 1-2 mg of glutathione-S-transferase (GST)-tagged versions of GCD7 and GCN3, from which the GST moiety can be cleaved in vitro by thrombin, and a His-tagged version of GCD2. We will attempt to form a trimeric complex with purified GCD2, GCN3 and GST-GCD7 fusion proteins and isolate the complex on a glutathione column. If we find that GCD2 must be present in the reactions in order to retain GCN3 on a glutathione column with GST-GCD7, this will be a good indication of specific complex formation between all three proteins. One important application for this assay will be to identify residues in GCD2 that are required for its interaction with GCD7 and GCN3. Initially, we will analyze the effects of deleting different segments of GCD2 on its ability to be specifically retained by the GST-GCD7 protein on a glutathione column. It should not be necessary to purify the various mutant GCD2 proteins from the bacterial extracts since their binding to the column can be assayed by immunoblot analysis. We should be able to study interactions between GCD2 and GCD7 by this approach even if the trimeric complex with GCN3 cannot be formed, since we have in vivo evidence for a GCD2-GCD7 complex. Using the His-tagged version of GCD2, we can also identify residues in GCD7 required for its interaction with GCD2, and it will be interesting to learn whether regions of homology between the two proteins mediate their physical interactions with one another.

If we succeed in reconstituting a GCD2-GCD7-GCN3 subcomplex, we will attempt to demonstrate binding of purified eIF-2 to this complex and analyze the effects of phosphorylating eIF-2 on the interaction. Wild-type eIF-2 will be purified according to our previously published procedures or as described below, and phosphorylated in vitro using recombinant PKR as already described. The eIF-2 will be applied to a nickel or glutathione column containing bound GCD2-GCD7-GCN3 trimeric complex and its retention on the column will be analyzed or by immunoblotting using antibodies against subunits of eIF-2. The K_D for the mammalian eIF-2/eIF-2B complex has been estimated between 0.1-1.0 nM (S. Kimball, personal communication). Thus, even if the K_D for interaction between eIF-2 and the GCD2-GCD7-GCN3 trimeric subcomplex is only 1 μ M, we should be able to detect it using this technique if we bind ca. 25 μ g of the reconstituted trimeric complex to ca. 50 μ L of resin. We will establish the specificity of an interaction between eIF-2 and the eIF-2B

subcomplex by showing that it depends on both GCD2 and GCD7 and that the affinity is higher when GCN3 is also present in the complex. We anticipate that phosphorylation of eIF-2 will increase its affinity for the trimeric complex and that introduction of one or more Gcn⁻ mutations into the GCD2, GCD7 or GCN3 proteins will either diminish the affinity for phosphorylated eIF-2 or increase the affinity for nonphosphorylated eIF-2. An increase in affinity will be detected by retention of the same fraction of applied eIF-2 in the presence of less trimeric complex bound to the column.

2. Studies on the regulation of mammalian eIF-2B by phosphorylated eIF-2.

There is a high degree of sequence similarity between GCD2, GCD7 and GCN3 and the corresponding subunits of mammalian eIF-2B. Interestingly, a number of the Gcn⁻ alleles we isolated in these subunits alter residues that are conserved between the yeast and mammalian proteins. We wish to determine whether these amino acids are involved in regulating mammalian eIF-2B by eIF-2(α P) in the same way that they function in yeast. To test this idea, we used site-directed mutagenesis to introduce the amino acids present in five Gcn⁻ alleles of *GCD2* at the corresponding positions of the rat eIF-2B δ cDNA. These mutant cDNAs are being inserted into different vectors designed for expression in mammalian cells. We will examine the effects of expressing these mutant subunits of eIF-2B on the response to eIF-2 phosphorylation in mammalian cells by the following two assays.

It has been established that certain expression vectors activate PKR in COS cells and this leads to greatly reduced expression of the DHFR gene contained on the same plasmid. This inhibition has been attributed to localized activation of PKR by double-stranded RNAs formed between DHFR mRNA and antisense RNAs made from the same plasmid. Overexpression of the non-phosphorylatable eIF-2 α -S51A protein from a different plasmid can rescue expression of DHFR from the PKR-sensitive vector. The overexpressed eIF-2 α -S51A protein replaces wild-type eIF-2 α in the trimeric eIF-2 complex and thereby reduces the amount of eIF-2 in the cell containing the phosphorylated form of the α subunit. We will determine whether overexpression of the δ subunit of rat eIF-2B bearing one or more Gcn⁻ mutations has the same effect on DHFR expression as that described for eIF-2 α (S51A). Each of the constructs containing wild-type or mutant rat eIF-2B δ cDNAs inserted in the vector pMTVA- will be co-transfected into COS-1 cells with plasmid pD61 encoding the DHFR reporter gene and expression of DHFR will be analyzed as described. As controls, the DHFR reporter plasmid will be co-transfected with pMTVA- derivatives encoding wild-type eIF-2 α or eIF-2 α -S51A, provided by R. Kaufman. If we find that expression of one or more mutant forms of eIF-2B δ leads to increased DHFR expression relative to that given by the wild-type δ subunit, this will imply that the mutation renders mammalian eIF-2B less sensitive to the inhibitory effects of eIF-2(α P), just as occurs in yeast. These experiments should also reveal the extent to which eIF-2B mediates the regulatory effects of eIF-2(α P) in mammalian cells.

In collaboration with Scot Kimball's lab the same mutant and wild-type rat eIF-2B δ cDNAs will be expressed from the pBK-RSV vector in rat cells and analyzed for their effects on eIF-2B nucleotide exchange activity in cell extracts. They have established procedures for increasing eIF-2 α phosphorylation by treating rat GH3 pituitary tumor cells with Brefeldin A and shown that eIF-2B in the whole cell extracts exhibits reduced nucleotide exchange activity for eIF-2-[³H]GDP binary complexes added as substrate. We will determine whether overexpression of any of the mutant forms of eIF-2B δ in the rat cells can rescue eIF-2B exchange activity. It is hoped that the various Gcn⁻ mutations will have similar relative effects on eIF-2B activity in COS-1 cells, as measured indirectly by rescue of DHFR expression, and in the rat cells as measured directly by assaying eIF-2B exchange activity.

C. Studies on GCD10 and the eIF-3 complex in GCN4 translational control and general translation initiation.

1. Domain structure of GCD10 and identification of its RNA binding site.

To begin our analysis of the role of RNA binding by GCD10 in eIF-3 function we will identify the minimal segment of GCD10 required for RNA binding in vitro and evaluate its importance for GCD10 function in vivo. We will construct nested deletions progressing from the N- and C-termini of GCD10 in a His-tagged allele tailored for expression in *E. coli* and for in vitro transcription/translation in rabbit reticulocytes. The mutant proteins will be tested for RNA binding by Northwestern and by binding to RNA immobilized on Separose. After identifying the minimal fragment of GCD10 required for RNA binding in vitro, we will analyze the in vivo effects of deletions in this domain. Assuming that the RRM-like sequence will be required for RNA binding in vitro, we will also examine effects of point mutations in the conserved RNP-1 and RNP-2 elements on RNA-binding in vitro and GCD10 function in vivo. For all these experiments, a myc epitope-tagged *GCD10* allele will be mutagenized and introduced into yeast in place of wild-type *GCD10* by plasmid shuffling. We will determine whether the *gcd10* mutations reduce growth rate or lead to derepression of *GCN4* (Gcd⁻ phenotype). If the mutations are lethal, they will be tested for a dominant slow-growth or Gcd⁻ phenotype. If a dominant-negative phenotype is observed, we will determine whether the mutant protein competes with wild-type GCD10 for incorporation into eIF-3 by immunoprecipitating with antibodies against PRT1 and analyzing the immune complexes with anti-myc antibodies. If so, mutagenesis of the dominant-negative allele will be carried out to isolate mutations that abolish this phenotype without reducing the level of the mutant protein. These mutations should identify residues in GCD10 required for stable complex formation with the other eIF-3 subunits. This interpretation will be tested by coimmunoprecipitation experiments. We will also rescue the existing *gcd10* chromosomal mutations by plasmid gap-repair and sequence the rescued alleles. These mutations may identify a functional domain of GCD10 that is distinct from those involved in RNA binding and eIF-3 subunit interactions but which is critical for repression of *GCN4* translation; eg. a domain that interacts with eIF-2. If so, we will begin in vitro mutagenesis of *GCD10* to isolate additional mutations affecting this domain by mutagenizing a high-copy plasmid bearing *GCD10* and selecting for growth on 5-FU plates (Gcd⁻ phenotype) in a strain containing wild-type chromosomal *GCD10*. By selecting for dominant Gcd⁻ mutations, we should isolate mutations that alter GCD10 function rather than simply reducing its expression or ability to interact with other components of eIF-3.

2. Genetic identification of proteins that interact with GCD10.

We are in the process of localizing and sequencing the remaining high copy suppressor of *gcd10* mutations to determine if it encodes an additional subunit of eIF-3. We also will begin to screen high copy plasmid genomic libraries for suppressors of a *prt1* temperature-sensitive mutation, and a P_{GAL1}-cDNA expression library for suppressors of both *gcd10* and *prt1* mutations. This suppressor analysis may identify proteins that specifically interact with these subunits of eIF-3 in addition to isolating the remaining subunits of the eIF-3 complex. If we do not obtain any new suppressor genes in these approaches, we will turn to the isolation of synthetic lethal mutations that render *gcd10* or *prt1* mutants inviable at a permissive or semi-permissive temperature.

John Hershey's lab at U.C.-Davis is obtaining peptide sequences from the 135, 33, 29 and 21 kDa subunits of yeast eIF-3 and has agreed to share this information with us. By this approach, they have already identified the 39 and 135 kDa subunits as sequenced ORFs of previously unknown function, encoded on Chromosome XIII, whereas none of our suppressors maps to XIII. Presumably, the structural genes for the remaining subunits of eIF-3 will be cloned either as one of our high copy suppressors or from the peptide sequences obtained in Hershey's lab. It is important to verify that all of the proteins that co-purify with eIF-3 in Hershey's purification scheme can be co-immunoprecipitated with the GCD10 and PRT1 proteins. Thus, we will construct epitope tagged alleles of the 39 and 135kDa subunits (and any others that are identified in the yeast database

from the peptide sequences) and replace the wild-type chromosomal alleles with the tagged versions. These strains will be used to determine whether the epitope-tagged proteins can be quantitatively immunoprecipitated with GCD10 and PRT1. These cloned genes can be used in the future to study interactions between GCD10 and other subunits of eIF-3 with the same techniques described above for studying GCD7-GCD2 interactions.

3. Analysis of *GCD13* and *GCD14*.

To learn more about the function of GCD14, we will test it for RNA binding activity using the assays described above for GCD10. We will also analyze existing temperature-sensitive *gcd14* mutations for polysome profiles indicative of general defects in translation initiation. We will construct an epitope-tagged functional allele of *GCD14*, or raise antibodies against a trpE fusion protein expressed in bacteria. Using antibodies against GCD14, we will investigate whether the protein is associated with polysomes, 40S or 60S subunits fractionated by velocity sedimentation on sucrose gradients, and also test for interactions with eIF-2·eIF-2B or eIF-3 by co-immunoprecipitation experiments. The isolation of dosage suppressors of *gcd14* mutations will also be initiated. We will screen the P_{GAL1}-cDNA expression library for a *GCD13* cDNA clone.

D. Regulation of protein kinase GCN2

1. Analysis of the N-terminal regulatory domain and potential autophosphorylation sites.

Ilse Barthelmess' group isolated a gene from *N. crassa* (*cpc-3*) encoding a GCN2 homologue (NcGCN2), with 40% and 32% sequence identity to the protein kinase and histidyl-tRNA synthetase-related domains of GCN2, respectively. Deletion of *cpc-3* impairs derepression of the *GCN4* homologue in *Neurospora*, called *cpc-1*. In collaboration with their lab, we are completing the sequence of an ca. 500-residue N-terminal domain of NcGCN2 and find that it shows significant similarity to the corresponding region in GCN2 (25% identity). In both proteins, this region contains a degenerate kinase catalytic domain of ca. 300 residues. In GCN2, there are consensus copies of kinase subdomains VII-XII in the correct order, whereas the NcGCN2 pseudo kinase domain (ΦPK) contains subdomains II, and VIII-XI. N-terminal to the ΦPK, both GCN2 and NcGCN2 contain an ca. 80-amino acid segment rich in positively and negatively charged residues. It seems likely that this charged segment is located on the surface of the protein and, could perhaps be a binding site for an effector protein.

Two instances of ΦPK domains have been reported in mammalian proteins. The JAK family of tyrosine protein kinases (that interact with cytokine receptors) contain a ΦPK domain of unknown function in addition to a conventional PK domain, whereas atrial natriuretic peptide (ANP) receptor contains a ΦPK domain N-terminal to a guanylate cyclase catalytic domain. For the latter, the ΦPK segment appears to be a negative regulatory domain that mediates stimulation of the cyclase activity by ATP and also provides a binding domain for a novel protein phosphatase that may be involved in desensitizing the the ANP receptor by dephosphorylation. By analogy with the ANP receptor, the ΦPK domain of GCN2 may play an important role in regulating kinase activity in response to amino acid or purine limitation, or under unknown conditions of starvation or stress, by binding another kinase or phosphatase that modifies the authentic kinase domain, as there is evidence that GCN2 is phosphorylated by another kinase *in vivo*. The ΦPK domain contains all the conserved residues in the C-terminal lobe located near the peptide substrate binding groove in PKA therefore, it might compete with the authentic kinase domain for binding eIF-2α. Alternatively, it might physically interact with the authentic kinase domain, or mediate dimer formation between two molecules of GCN2.

We will take several genetic approaches to address the importance of the ΦPK domain. In one approach, we will use site-directed mutagenesis to substitute conserved amino acids in ΦPK subdomains VII-XII with alanines. In addition, we will construct chimeras that replace the yeast ΦPK domain with that from the *Neurospora* homologue. If this chimera is functional, it will imply that the exact set of invariant residues left intact is not critically important for ΦPK function. We will also generate small in-frame deletions throughout

the region (using two-codon *SacI* site insertions made previously, and mutagenize the Φ PK domain at random sites in the *E. coli* mutator strain XL1-Red (Stratagene), and screen the mutant plasmids for reduced GCN2 function (Gcn⁻ phenotype) or constitutive activation of GCN2 (Gcd⁻ phenotype) using amino acid analogues as indicators of *GCN4* expression. This random mutagenesis should identify residues in the Φ PK that are not conserved among protein kinases, but which are involved in the regulatory function of the domain, either as binding sites for other regulatory proteins or as points of contact between the pseudo and authentic kinase domains or between protomers in a GCN2 homodimer. We will also make deletions in the 80-residue N-terminal highly charged segment and, depending on their phenotypes, introduce clustered alanine substitutions throughout this segment of GCN2. In addition, we will replace it with the corresponding segment from *Neurospora* NcGCN2 which, although highly charged, shows little primary sequence similarity to the yeast segment.

Another very interesting target for mutagenesis that could be involved in the function of the Φ PK domain is a segment between subdomains VII and VIII of the authentic PK domain of GCN2. The activity of several protein kinases, including PKA, MAPK, MAPKK and CDK requires phosphorylation or autophosphorylation of one or two closely spaced Ser or Thr residues in this segment. In the case of CDK, this phosphorylation is thought to alter the conformation of a loop that protrudes into the active site and prevents proper positioning of the protein substrate. It can also promote proper interactions between the N-terminal and C-terminal lobes of the kinase domain, as in the case of PKA and MAPK. GCN2 contains a larger than usual segment separating subdomains VII and VIII, containing a total of 7 serine and threonine residues. One or more of these residues could be critical sites of autophosphorylation or phosphorylation by another kinase, or sites of dephosphorylation by a protein phosphatase. We have mutated one of these residues, Thr-813, to Gly and observed a leaky Gcn⁻ phenotype. We will mutate additional Thr and Ser residues in this segment to Ala, singly and in combination with an Ala substitution at Thr-813, and analyze the effects on *GCN4* expression in vivo. If one or more of these residues is found to be critical for GCN2 function in vivo, we will determine whether function can be restored by substitutions with Asp or Glu, which may functionally substitute for phosphoserine or phosphothreonine. In addition, we will determine whether the mutations reduce autophosphorylation of GCN2 in our standard immune complex assays, or its level of phosphorylation in vivo by labeling cells with radioactive phosphate, immunoprecipitating GCN2 from whole cell extracts, and analyzing the labeled protein by SDS-PAGE. If there are multiple phosphorylation sites, it will be necessary to analyze the tryptic phosphopeptides produced in vitro and in vivo by mutant and wild-type proteins. If the critical site is for autophosphorylation, it will be necessary to show that substitution of the site with Ser (if the wild-type residue is Thr) or Thr (if the wild-type residue is Ser) leads to the production of a novel phosphopeptide. If autophosphorylation occurs in trans following dimerization, we may also succeed in showing that wild-type GCN2 can phosphorylate the critical residue using a K559V defective kinase as substrate.

2. Analysis of GCN1/GCN20 interactions.

The importance of the EF-3-related domain of GCN1 for protein-protein interactions with GCN20 suggests that the positive regulatory function of GCN1/GCN20 is related to the role of EF-3 in translation elongation. To pursue this possibility further, we would like to determine whether residues in GCN1 that are conserved with EF-3 are of critical importance for *GCN1* function. We also wish to identify specific amino acids that are required for protein-protein interactions with GCN20, as these may illuminate the role of GCN20 in the putative EF-3-like function of GCN1. GCN1 encodes a very large protein of 2,672 amino acids, and in addition to mutational analysis of the EF-3-related domain (residues 1329-2140), we would like to determine whether other segments of the protein are important for its regulatory function. We will begin a mutational analysis of GCN1 by testing internal deletions of ca. 300-amino acid segments, constructed using naturally occurring and engineered restriction sites for: (1) loss of *GCN1* function, revealed by sensitivity to 3-AT; and (2) constitutively activated *GCN1* function, leading to 5-FT resistance. This approach may allow us to eliminate relatively large

segments of GCN1 that may not be required for its regulatory function. In a second approach, we have used the *E. coli* mutator strain XL-1 Red (Stratagene) to carry out random mutagenesis of *GCN1*, and we are subcloning segments of 1-2 kb from this plasmid pool into an unmutagenized *GCN1* plasmid pool. Screening of these pools will be restricted to those segments found to be important for GCN1 function by the internal deletion analysis, beginning with the segment corresponding to the EF-3-related region of GCN1. Mutant alleles will be subjected to Western analysis and those which produce full-length GCN1 at wild-type levels will be sequenced to identify the mutation. The mutant proteins will be tested for protein-protein interactions with GCN20 using the co-immunoprecipitation assays described above. Any *gcn1* mutations that reduce co-immunoprecipitation of the two proteins will be examined for the ability to be suppressed by overexpression of GCN20 from a high copy plasmid. If an in vitro assay for GCN2 kinase activity showing GCN1-dependence is available at this time, the *GCN1* mutations will be analyzed for effects on GCN2 kinase function in vitro. Any *GCN1^c* alleles we obtain will be tested to determine whether they are dependent on *GCN20* for their derepressed phenotype.

3. Genetic screening for novel proteins that interact with GCN2.

At present, we have not detected a physical interaction between the GCN1/GCN20 complex and GCN2 and little is known about how these proteins regulate GCN2 kinase activity by uncharged tRNA. The location of GCN2 on the ribosome and possible roles of specific ribosomal proteins in this regulation are also unknown, as is the mechanism for coupling GCN2 kinase activity to purine levels. It is unlikely that the *gcn* class of mutations has been saturated, and there may be additional proteins besides GCN1 and GCN20 involved in controlling GCN2 activity. For these reasons, we will pursue several genetic approaches aimed at the identification of proteins that interact with GCN2. In one approach, we will use yeast two-hybrid constructs containing different segments of GCN2 to screen the Elledge cDNA library and the Clontech genomic library. We will begin with a bait construct containing the Φ PK domain of GCN2 and proceed as necessary to one containing the HisRS-like domain, following the same plan outlined above for analyzing two-hybrid interactions with GCN1 bait plasmids.

In a second approach, we will attempt to saturate the *gcn* class of mutations by random insertional mutagenesis of the yeast genome using the insertion library constructed in Mike Snyder's lab. This is a plasmid library of genomic fragments containing Tn3::lacZ::LEU2 transposons inserted at random sites. The mutagenized yeast fragments are removed from the plasmid vector by digesting with NotI and then used to make chromosomal insertions by transforming a yeast strain to Leu⁺. Using this library, we will mutagenize a haploid strain that is wild-type for all known *GCN* and *GCD* genes (except for *GCN4*) and contains an integrated *HIS3-gus* fusion under the control of GCN4 and an integrated *GCN4-lacZ* fusion. The strain will carry one of our mutant alleles of *GCN4* which contains only the N-terminal activation domain. This *GCN4* allele confers high-level activation of GCN4-regulated genes under our standard starvation conditions but is defective relative to wild-type *GCN4* under more severe starvation conditions. By incorporating this attenuated allele in place of wild-type *GCN4*, we hope to increase the sensitivity of our screen for *gcn* mutations. (In addition, we may identify mutations in nonessential transcriptional co-activator proteins, such as *GCN5*, that impair transcriptional activation mediated by the N-terminal activation domain; see below). Transformants containing insertions in different loci will be screened for increased sensitivity to 3-AT, sulfometuron (inhibits Ile and Val synthesis), and azaadenine (azaA) (inhibits purine biosynthesis), indicating a Gcn⁻ phenotype. Transformants showing sensitivity to 3-AT and SM or to azaA will be subjected to quantitative assays of the two integrated fusions after growth in the presence of 3-AT or azaA and ranked according to the magnitude of their derepression defects. (Note that 1/6 of the chromosomal insertions will fuse *lacZ* in-frame to the disrupted gene. For this class of insertion mutants, we will conduct Western analysis of *GCN4* expression using antibodies against GCN4.) Any mutants showing reduced *GCN4* expression under starvation conditions will be analyzed genetically to determine

whether this phenotype results from a single insertion. Mutants showing reduced *HIS3-gus* expression but normal *GCN4* expression will be set aside for studies on *GCN4* transcriptional co-activators (see below). For all others, we will isolate genomic sequences adjacent to the insertion site using the plasmid integration-excision scheme of Burns et al. and determine the sequence of 100-200 nt. If the sequence does not correspond to a known *GCN* gene or match an entry in the yeast database, we will determine the map location by hybridization to the Riles-Olson lambda library and clone the wild-type gene from the appropriate lambda or cosmid clone purchased from ATCC. The gene will be localized by complementation testing, completely sequenced, and a chromosomal deletion will be constructed to determine the null phenotype with respect to *GCN4* expression and eIF-2 α phosphorylation.

E. Molecular mechanism of transcriptional activation by GCN4

1. Biochemical analysis of potential interactions between GCN4 and transcriptional mediators.

Our mutational analysis of *GCN4* has identified two functionally distinguishable activation domains that each contain multiple clusters of 2-3 aromatic or bulky hydrophobic amino acids that are critically required for the function of that domain. Five such clusters have been identified in the NTAD and 2 are located in the CAAD. Some or all of these residues could mediate important interactions between *GCN4* and its target proteins in the transcriptional machinery. Genetic analysis has implicated the *ADA2*, *ADA3* and *GCN5* proteins as mediators of transcriptional activation by *GCN4* and several other acidic activators; however, at present, there is no evidence that *GCN4* physically interacts with these proteins or with any of the general transcription factors. We plan to use the point mutations we generated in the activation domains of *GCN4* to identify target proteins in the transcription machinery and to identify the domains in these proteins that contact *GCN4*.

We are testing *GCN4*-GST fusion proteins for interactions with candidate proteins synthesized by in vitro transcription/translation, focusing initially on the TATA-binding protein TBP. A panel of *GCN4*-GST fusion proteins are being overexpressed in *E. coli*, including full-length *GCN4*, deletion derivatives lacking the NTAD, CAAD, or both, and several full-length constructs containing multiple point mutations in both the NTAD and CAAD. Bacterial extracts containing these proteins are mixed with aliquots of in vitro translation extracts containing ³⁵S-methionine-labeled yeast TBP, after which we determine the amount of radiolabeled TBP that is retained with the GST-*GCN4* proteins on glutathione-agarose beads (GST pull-down assay). We have also carried out reciprocal reactions in which ³⁵S-methionine-labeled *GCN4* proteins synthesized in vitro are tested for interactions with a GST-TBP fusion protein. In control experiments, we have reproduced the interaction between VP16 and TBP that depends on the C-terminal activation domain of VP16. In addition, we observed *GCN4*-*GCN4* homodimer formation. Although we detected an interaction between TBP and *GCN4*, it appeared to be relatively weak compared to the TBP-VP16 interaction. In addition, the DNA binding-dimerization domain at the C-terminus of *GCN4* seemed to be both necessary and sufficient for the in vitro interaction with TBP. Consequently, we will not pursue this interaction for the time being. We also used the GST pull-down assay to explore interactions between yeast TFIIB and *GCN4*, but thus far have not observed any interaction between them. In future experiments we will to examine possible interactions between *GCN4* and the individual *ADA2*, *ADA3* and *GCN5* proteins synthesized by in vitro transcription/translation, as described above.

It is possible that *GCN4* can interact with the *ADA2*, *ADA3* and *GCN5* proteins only when the latter are present in a complex. To explore this possibility, we will add GST-tagged mutant and wild-type *GCN4* proteins expressed in *E. coli* to whole cell yeast extracts prepared from strains expressing either untagged *ADA2* or a functional HA-tagged form of *ADA2* (provided by L. Guarente). We will then determine what fraction of the *GCN4* proteins can be co-immunoprecipitated with HA-*ADA2* and what fraction of HA-*ADA2* can be co-purified with *GCN4* on a glutathione-agarose column. Similar experiments have been used to demonstrate that GAL4-VP16 and native GAL4 can interact with *ADA2* or with the mediator complex.

We will also examine whether components of the holoenzyme mediator complex, purified from yeast as described by Kim et al, can be recovered with GCN4-GST proteins on glutathione agarose. This last experiment will be done in collaboration with Christoph Hengartner in Rick Young's who has provided us with partially purified holoenzyme and antibodies against several of the SRB protein components of the mediator complex. In addition, Tony Weil has supplied us with a partially purified preparation of a TFIID-like complex from yeast, consisting of TBP and associated factors (A. Weil, personal communication). We will probe for interactions between these proteins and GST-GCN4 proteins using glutathione agarose chromatography and antibodies against some of the TFIID protein components provided by Weil's group. If we detect an interaction between GCN4 and the mediator or TFIID complex that is abolished by one or more mutations in the GCN4 activation domain, we will follow up with GST-GCN4 pull-down experiments using individual subunits of the appropriate complex. Young's and Weil's groups are in the process of purifying these proteins after overexpressing them in insect cells from baculovirus vectors, and they have indicated their willingness to collaborate with us in testing the purified proteins for specific binding to GST-GCN4 proteins in vitro.

2. Genetic screens for transcriptional factors that interact with GCN4.

a. Suppressors of a mutant GCN4 protein that is lethal when overexpressed.

We have set up several genetic screens aimed at identifying functionally important interactions between GCN4 and components of the transcriptional machinery. One approach already underway is to isolate chromosomal mutations that overcome the dominant lethal effect of overexpressing a mutant GCN4 protein which lacks the CAAD but retains the NTAD, and is produced at ca. 4-fold higher levels than wild-type GCN4. Because the toxic effect of this construct is dependent on both a functional DNA binding domain and activation determinants in the NTAD, we presume that the mutant GCN4 protein sequesters one or more essential transcription factors. We reasoned that it might be possible to isolate mutations that eliminate a nonessential mediator that is required for sequestering the essential transcription factor by the mutant GCN4 protein, as this approach was used previously to isolate mutations in *ADA2* that overcome the toxicity of an overexpressed GAL4-VP16 protein in yeast. Since we found that the toxicity of overexpressing the mutant GCN4 protein is not suppressed by deletion of *ADA2*, we hope to identify a distinct mediator that interacts with the NTAD of GCN4. It might also be possible to isolate mutations in a general transcription factor that reduces its affinity for GCN4 without eliminating essential interactions with other activators.

We have isolated spontaneous revertants of the lethal phenotype conferred by overexpressing the mutant GCN4 protein from an uORF-less construct carried on a single-copy *URA3* plasmid. After removing the toxic *GCN4* allele from the revertants by 5-FOA counter-selection, we identified strains which exhibited 3-AT-sensitivity, indicating a possible defect in the ability of the resident wild-type GCN4 to activate transcription. Meiotic analysis was used to identify those revertants in which 3-AT sensitivity (in the presence of wild-type *GCN4*) and suppression of the lethal *GCN4* mutant allele were conferred by a single mutation. Several of these mutants also displayed a slow-growth phenotype on rich medium (Slg⁻) that co-segregated with the other phenotypes. For most of these Slg⁻ revertants, this phenotype could be partially complemented by introducing a single-copy plasmid bearing *SPT13*, encoding yeast TBP. Accordingly, the endogenous *SPT13* gene was isolated from these revertants by gap repair and was found to be indistinguishable from wild-type *SPT13* in complementing the Slg⁻ phenotype of our revertants. Hence, none of these revertants appear to contain suppressor mutations in *SPT13*. The fact that an extra copy of *SPT13* partially complemented their Slg⁻ phenotype does suggest, however, that they contain mutations affecting a general transcription factor. We have verified that the suppressor mutations do not reduce the steady-state level of the GCN4 protein, and we are attempting to clone the wild-type alleles for three of the suppressor mutations with particularly strong Slg⁻ phenotypes.

In a related approach, we will screen a high copy-number plasmid library for genes that when overexpressed can overcome the lethality of overexpressing the mutant GCN4 protein described above. If this *GCN4* construct is lethal because a co-activator or general factor is being sequestered by the mutant GCN4 protein, then overexpressing the sequestered factor should relieve the toxicity. Even if the sequestered factor is a multimeric complex, this approach could succeed because overexpressing only one subunit of the complex might be enough to sequester the toxic GCN4 protein and prevent it from interacting with the native co-activator complex. A high copy-number *URA3* plasmid library will be introduced into a wild-type *GCN4* yeast strain and a pool of ca. 10^6 transformants will be isolated and transformed with the toxic *GCN4* allele on a single copy-number *LEU2* plasmid. Transformants that arise will be tested for the ability to lose the library plasmid and grow on 5-FOA. Failure to grow on 5-FOA except after loss of the *LEU2* plasmid will indicate the presence of a dosage suppressor of the lethal *GCN4* construct. Transformants of this type will be analyzed to determine whether the suppressor plasmid lowers the expression of the mutant GCN4 protein, and then tested for exacerbation or suppression of the Gcn⁻ phenotype of a panel of leaky *gcn4* alleles bearing different mutations in the activation domain. Exacerbation of the Gcn⁻ phenotype could occur if the encoded protein is a component of a heteromeric co-activator complex that can compete with native co-activator for interaction with defective GCN4 proteins. Suppression of the Gcn⁻ phenotype could occur if the overexpressed protein is a monomeric co-activator whose interaction with GCN4 is impaired by one of the *gcn4* mutations being tested. We will also determine the null phenotype of the suppressor gene. If a deletion strain is viable, we will examine it for Gcn⁻ phenotypes (3-AT-sensitivity and defective derepression of *HIS-lacZ* fusions) and for defects in activation by GAL4 and HAP4 using the appropriate *lacZ* reporter constructs dependent on these activator proteins. We will epitope-tag the gene with the *myc* epitope and use anti-myc antibodies to probe for co-immunoprecipitation of the suppressor gene product with components of the ADA2/ADA3/GCN5 complex, or with the mediator or TFIID complexes described above. The suppressor protein will also be tested for direct interactions with GCN4 protein in vitro using the GST pull-down assay described above.

b. Suppressors of a leaky *GCN4* allele with point mutations in the activation domains.

In an independent genetic screen, we will isolate chromosomal suppressors of the 3-AT-sensitive phenotype of the *gcn4-1843* allele, which contains alanine substitutions in two of the activation modules in the CAAD (at Met-107, Tyr-110, Leu-113 and Trp120, Leu-123, and Phe-124) and in one activation module of the NTAD (Phe-97 and Phe-98). Because it contains mutations in both the CAAD and NTAD, this allele confers very weak activation in vivo despite an ca. 4-fold higher than wild-type level of protein. Assuming that these mutations weaken interactions between GCN4 and one or more co-activator proteins or basic transcription factors, we anticipate that chromosomal mutations can be isolated in these factors that will restore their interaction with mutant GCN4. By testing these mutations for the ability to suppress other *gcn4* alleles that contain a different set of point mutations, we may identify suppressors that specifically restore interactions with particular activation modules of GCN4. We may also isolate mutations in other components of the transcription machinery which bypass the requirements for those activation modules which are lacking in the *gcn4* allele. It is conceivable that certain transcription factors that are normally recruited or activated by GCN4 could be mutated to mimic the effect on them that wild-type GCN4 normally exerts. Even if they map in known proteins, such bypass mutations would highlight the importance of these proteins in transcriptional activation by GCN4.

We have constructed a pair of isogenic strains containing *gcn4-1843* integrated on a *URA3* plasmid at *ura3-52*, plus a *GCN4-lacZ* fusion at *TRP1* and a *HIS3-GUS* construct at *LEU2*. We are collecting 3-AT-resistant revertants in one of these strains and cross-screening them for resistance to sulfometuron-methyl (SM) and 5-methyltryptophan (5-MT) (that inhibit isoleucine-valine and tryptophan biosynthesis, respectively), and for resistance to cyclohexamide (to screen out multiple drug resistance). Revertants that are resistant to 3-AT, SM

and 5-MT, but sensitive to cyclohexamide, will be assayed for expression of the *HIS3-GUS* and *GCN4-lacZ* reporters. Those showing increased *HIS3-GUS* expression but unchanged *GCN4-lacZ* expression will be analyzed to determine whether they contain a back-mutation at *GCN4* or an unlinked suppressor. This will be done by evicting the *gcn4-1843* allele on the integrating *URA3* plasmid on 5-FOA medium, followed by transformation with a single copy plasmid bearing *gcn4-1843*. Revertants containing trans-acting suppressors will revert to a 3-AT-sensitive phenotype upon eviction of the *gcn4-1843* allele and re-acquire 3-AT-resistance upon reintroduction of *gcn4-1843* on a plasmid. Revertants that retain the 3-AT^r phenotype after eviction of the *gcn4-1843* allele will be set aside for the present, because they contain GCN4-independent suppressors. All revertants that contain unlinked GCN4-dependent suppressors will be crossed to the isogenic strain containing the *gcn4-1843* allele and subjected to meiotic analysis. Revertants found to contain single suppressor mutations responsible for their 3-AT-resistant phenotypes will be tested for allele-specific suppression by transforming the versions of the revertant strains which lack the integrated *gcn4-1843* with plasmid-borne *gcn4* alleles containing different point mutations in the activation domain. Suppressors that fail to suppress, or which even exacerbate, the 3-AT-sensitive phenotype of other leaky *gcn4* alleles will be particularly interesting. These may contain a mutation in a co-activator protein that restores its interaction with an altered GCN4 activation module but that weakens its interaction with the wild-type version of that module. Genetic analysis will be used to assign suppressors to complementation and linkage groups and the wild-type genes will be cloned by complementing either the suppressor phenotype, or a suitable secondary phenotype associated with the suppressor mutations.

The same strain just described containing *gcn4-1843* is also being used to screen high copy plasmid yeast libraries for dosage suppressors. A mediator protein or basic transcription factor that functions as an individual polypeptide might suppress a GCN4 activation defect when overproduced. In addition, if GCN4 is phosphorylated as a means of stimulating its activation function, we might identify a protein kinase, an activator of a kinase, or an inhibitor of a phosphatase as a dosage suppressor of a defective *gcn4* allele. The virtue of this approach is that a plasmid library can be screened rapidly by direct selection for 3-AT-resistant transformants, and plasmid-associated suppressors can be readily distinguished from chromosomal suppressors by determining whether the suppressor phenotype disappears when the plasmid is lost from the strain. The same genetic criteria listed above for chromosomal suppressors will be applied to evaluate high copy-number plasmid suppressors.

F. Regulation of purine nucleotide biosynthesis

Our interpretations of the mutational analysis of the *ADE5,7* promoter will be evaluated by conducting gel retardation assays on various mutant and wild-type DNA fragments using BAS1, BAS2 and ABF1 proteins expressed in *E. coli*, or produced by in vitro transcription/translation, following published procedures. From these experiments we hope to learn whether the requirement for particular nucleotide positions for transcription and adenine repression in vivo can be accounted for by their effects on binding BAS1, BAS2, or ABF1 in vitro. They should also reveal the identify of the bindingsite(s) for BAS2 at *ADE5,7* and indicate whether the binding of one protein is dependent on the other two. We will continue with our analysis of protein fusions to identify the adenine-repressible activity of BAS1 or BAS2. In addition, we will proceed with our genetic selections for trans-acting mutations that abolish adenine-mediated repression of transcription to identify additional factors that might be involved in regulating BAS1 or BAS2 by adenine and to delineate the regulatory domains in BAS1 or BAS2, all as described above in the progress report.

Publications:

Drysdale CM, Duenas E, Jackson BM, Reusser U, Braus GH, Hinnebusch AG. Transcriptional activator GCN4 contains multiple activation domains that are critically dependent on hydrophobic amino acids. *Mol Cell Biol*, 1995;15:1220-33.

Flowers KM, Kimball SR, Feldhoff RC, Hinnebusch AG, Jefferson LS. Molecular cloning and characterization of cDNA encoding the α -subunit of the rat protein synthesis initiation factor eIF-2B. *Proc Natl Acad Sci USA*, 1995;92:4274-78.

Garcia-Barrio MT, Naranda T, Vazquez de Aldana CR, Cuesta R, Hinnebusch AG, Hershey JWB, Tamame M. GCD10, a translational repressor of GCN4, is the RNA-binding subunit of eukaryotic translation initiation factor-3. *Genes Dev*, 1995;9:1781-96.

Romano PR, Green SR, Barber GN, Mathews MB, Hinnebusch AG. Structural requirements for double-stranded RNA binding, dimerization, and activation of the human eIF-2 α kinase DAI in Saccharomyces cerevisiae. *Mol Cell Biol*, 1995;15:365-78.

Vazquez de Aldana C, Wek RC, Segundo PS, Truesdell AG, Hinnebusch AG. Multicopy tRNA genes functionally suppress mutations in yeast eIF-2 α kinase GCN2: Evidence for separate pathways coupling GCN4 expression to uncharged tRNA. *Mol Cell Biol* 1994;14:7920-32.

Vazquez de Aldana CR, Marton MJ, Hinnebusch AG. GCN20, a novel ABC protein, and GCN1 reside in a complex that mediates activation of the eIF-2 α kinase GCN2 in amino acid-starved cells. *EMBO J*, 1995;14:3184-99.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01005-08 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Regulation of Cellular Proliferation and Diversity in *Drosophila*

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	J. Kennison, Ph.D.	Senior Staff Fellow	LMG:NICHD
Others:	M. Vásquez, Ph.D.	Visiting Fellow	LMG:NICHD
	S. Stoler, Ph.D.	NRC Biotech. Fellow	LMG:NICHD
	A. Felsenfeld, Ph.D.	Guest Researcher	LMG:NICHD
	B. Brizuela	Bio. Lab. Tech	LMG:NICHD
	T. Brown	Student IRTA	LMG:NICHD
	A. Hobson	Student IRTA	LMG:NICHD
	R. Scott	Summer IRTA	LMG:NICHD

COOPERATING UNITS (if any)

Dr. J. Tamkun, Department of Biology, University of California, Santa Cruz
Dr. D.-H. Huang, Institute of Molecular Biology, Academia Sinica, Taiwan

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Section on Developmental Biology

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

4.0

PROFESSIONAL:

2.6

OTHER:

1.4

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Genes involved in the determination and differentiation of adult tissues in *Drosophila melanogaster* have been identified by genetic interactions with homeotic and segmentation genes already known to be required in these tissues. Many new genes required for transcriptional activation of developmental genes have been identified. One of these new genes, *brahma*, encodes a large nuclear protein conserved from yeast to man. *brahma* mutations have in turn been used to identify other interacting genes required for transcriptional activation. The *osa* gene, which shows allele-specific interactions with *brahma* mutations, has been cloned and encodes a single 3 kb transcript. We have shown that maternal expression of both *brahma* and *osa* is essential for early embryogenesis, while zygotic expression of both is essential late in embryogenesis and during larval growth. We have also identified a third gene [the *Pearl* gene] that may encode a protein that interacts with both *brahma* and *osa* proteins.

We have isolated and characterized a cis-regulatory mutation in the *Drosophila hedgehog* gene. The *hedgehog* gene encodes a secreted protein that is involved in cell-cell signalling during both embryonic segmentation and adult patterning. Our cis-regulatory *hedgehog* mutation derepresses transcription in larval cells that normally do not express *hedgehog* proteins, and the misexpression of *hedgehog* proteins causes neighboring cells to alter their developmental fates. We have used the phenotypes caused by *hedgehog* misexpression to isolate interacting mutations. These interacting mutations identify genes that are required for *hedgehog* expression or function. We have mapped twenty-four of these mutations to seven genes that are also required for homeotic gene transcription. These seven genes include the *brahma*, *osa*, *trithorax*, and *moira* genes that we have extensively characterized.

Project Description

Objectives:

During the development of a multicellular organism, cells within the organism acquire specialized functions in reproducible spatial and temporal patterns. These patterns result from both the synthesis of tissue-specific proteins and the differential activation or interaction of pre-existing gene products. All aspects of differentiation must be integrated to form these patterns. The regulation and integration of information depends on intercellular communication within the early embryo, and upon intracellular signalling subsequent to cell-cell interactions.

Our goal continues to be the understanding of the molecular mechanisms that operate to control development. A key step toward this goal is to identify all of the relevant genes, something that is currently only feasible through a combined molecular and genetic approach. The determination of body segments in *Drosophila melanogaster* provides an excellent system for the investigation of the genetic basis of pattern formation. The discovery of homeotic and segmentation regulatory genes, major classes of genes that affect segmental differentiation in *Drosophila* and are widely conserved among all eukaryotes, has in the past been largely accidental, but the genetic screens that we have performed have been designed to make the search for regulatory genes more systematic. Because of a desire both to identify additional members of the gene networks, and to focus on important regulatory interactions between these genes, screens were designed to identify new loci that are likely to be involved in regulatory interactions with other genes required for segment determination. We have identified a large number of new genes required for proper segment formation, and are characterizing these genes to determine their individual functions in development.

We are characterizing the regulatory steps defective in the genes identified in our genetic screens. The majority appear to encode transcriptional regulators of multiple homeotic and segmentation genes. We have also isolated and characterized a regulatory mutation in the *hedgehog* segmentation gene. The ectopic transcription of *hedgehog* in this mutation causes pattern defects in adult tissues. We are able to analyze the effects of our new regulatory mutations on *hedgehog* expression more easily because the normal pattern of *hedgehog* expression is simpler than the expression patterns of the homeotic genes. We are using a combination of molecular and genetic techniques to determine the function(s) of several new regulators. Because of the presence of one of these new transcriptional regulators in one of the nuclear extracts required for *in vitro* transcription (the TFIID extract), we are also isolating mutations in a gene that encodes a component of the basal transcription factor present in this extract.

Methods Employed:

Standard methods are used for the culture of *Drosophila*. The alkylating agent, ethyl methanesulfonate, gamma-ray exposure, and P-element-transposon mobilization are used in the induction and recovery of mutations. The mutations are mapped by meiotic recombination using standard genetic marker mutations or by *in situ* hybridization of labelled nucleic-acid probes to giant polytene chromosomes of the larval salivary gland. Recombinant DNA techniques are employed to isolate and analyze DNA segments from the *Drosophila* genome. Mutations are mapped within such isolated DNA segments by Southern blotting of DNA from mutant stocks or by DNA sequence analysis. Transcription mapping of chromosome regions is done by Northern blotting and by the analysis of cloned cDNA. Sequence analysis is by enzymatic methods. P-element-mediated mutagenesis and transformation follow the methods of Engels and Spradling and Rubin. Cloned cDNA segments are expressed into protein by the

use of fusion vectors with β -galactosidase and antisera are raised by standard methods. Embryos are fixed and stained with antibodies to various *Drosophila* proteins by the methods of Karr and Carroll or hybridized to labelled nucleic acid probes by the method of Tautz. Larval imaginal discs are fixed and stained with antibodies to *Drosophila* proteins by the methods of Pattatucci and Kaufman. Analysis of germline function follows the methods of Chou, Noll, and Perrimon.

Major Findings:

As part of our effort to characterize the function of the genes that we have identified as important in regulating expression of homeotic genes, we have also characterized a gain-of-function mutation in the *hedgehog* segmentation gene. The *hedgehog* protein is secreted and appears to be important in cell-cell signaling during development. We have shown that misexpression of the *hedgehog* protein in larval tissues induces neighboring cells to express the *dpp* gene, which encodes a TGF β -related growth factor. Misexpression of *dpp* then causes additional tissue growth and altered patterning. We have isolated over 60 dominant mutations that prevent the developmental consequences of *hedgehog* misexpression in this gain-of-function mutation. While we continue to isolate and characterize mutations that affect expression of either homeotic genes or *hedgehog*, a large part of our efforts are directed towards understanding the functions of the genes that we have already identified. We have mapped forty-nine of the new mutations that effect *hedgehog* expression. We have mapped twenty-five of these mutations to seventeen genes that we have not further characterized. The remaining twenty-four mutations (about half) map to seven genes that we had previously identified because they are required for homeotic gene transcription. We have focused much of our effort on these seven genes [the genes *trithorax*, *brahma*, *osa*, *moira*, *kohtalo*, *verthandi*, and *skuld*]. All seven genes are required for both *hedgehog* and homeotic gene transcription. Using a variety of methods, we have isolated 104 mutations in these seven genes and have mapped six of the genes to small chromosomal deficiencies. Genetic interactions between these seven genes suggest that they encode proteins that physically interact in the transcriptional activation of target genes.

We have done the most extensive analysis for the *brahma* gene. The *brahma* gene encodes a nuclear protein that is conserved from yeast to man. This protein is found in large protein complexes that include at least nine other polypeptides. The yeast homologue, SNF2/SWI2, is a DNA-dependent ATPase that is required for transcriptional activation of multiple target genes. It has been shown to function *in vitro* to enhance binding of specific transcriptional activators to their target DNA sequences when nucleosomes are present. There are at least two different *brahma* homologues in mammals, and these are also found in large protein complexes. We have found extensive interallelic complementation among *brahma* mutations, suggesting that there are at least two *brahma* polypeptides per protein complex. The whole protein complex may function as a dimer, since the size determined by gel filtration is more than twice the size expected from the molecular weights of the individual subunits.

Because of the evidence that *brahma* homologues in yeast and man function in large protein complexes, we began searching for other *Drosophila* genes that show genetic interactions with *brahma* mutations. We have isolated mutations in at least five new genes, but the most striking genetic interactions have been found with *osa* mutations. *osa* is one of the other genes that we first identified because it is required for transcription of *hedgehog* and the homeotic genes. Not only do *brahma* and *osa* mutations show genetic interactions, but they share many phenotypes. We have shown that maternal *brahma* and *osa* transcripts are required for early embryogenesis and that the zygotic transcripts are required for late embryogenesis. Both appear to be essential for imaginal cell differentiation in all tissues examined.

By transposon-tagging, we have cloned the *osa* gene and shown that it encodes a 3 kb transcript. We have isolated several cDNA clones for this 3 kb transcript and have partially sequenced the longest cDNA clone. We have also mapped another of the new *brahma* interacting mutations [the *Pearl* mutation] to a small genomic region that includes the *Drosophila* homologue of a yeast gene that genetically interacts with the yeast *brahma* homologue.

We have also continued our characterization of the requirements for *moira* gene function during development. We have shown that maternal expression of *moira* is required for normal oogenesis, and that zygotic expression is only required late in embryogenesis and in the larval tissues that form the head, thorax, and genital structures of the adult. Loss of *moira* function in the thorax during larval growth suggests that *moira* is required for proper functioning of the *Ubx* and *engrailed* genes in these tissues. Using a transgene that expresses the bacterial β -galactosidase protein under the control of *engrailed* cis-regulatory elements, we have shown that *moira* is required for transcription of *engrailed* in the cells that form the posterior adult wing.

We have also begun both developmental and molecular analyses of the *l(3)87Ca* gene, which is one of the genes required for homeotic gene expression that is not required for *hedgehog* gene expression. We have transposon-tagged the gene and isolated genomic DNA clones that include the site of the transposon insertion. We have identified at least two transcripts from this region of the genome, but have not yet determined whether either or both are products of the *l(3)87Ca* gene. Preliminary evidence suggests that the gene is essential for both oogenesis, and in somatic cell differentiation.

Significance to Biomedical Research and the Program of the Institute:

Understanding the genetic control of development and its molecular basis is a major aim of the Institute research program. Our work is aimed at understanding the network of genes involved in cellular determination. One aspect of the work, the characterization of the effects on development of mutations in genes involved in transcriptional regulation, should prove invaluable in understanding how cells acquire different fates during embryonic growth. The wealth of genetic material available in *Drosophila* from the past eighty years of research coupled with recent advances in the techniques for molecular analyses have given *Drosophila* many advantages among organisms for such a study. The finding of many similarities between *Drosophila* and mammalian genes has strengthened the belief that studies in *Drosophila* will lead to a greater understanding of basic problems in development in other animals.

Proposed Course of Project:

We propose to continue our developmental and molecular characterizations of the genes that are required for homeotic gene transcription. We will finish determining the sequence of the *osa* cDNA clones. The cDNA clones will then be used to generate antisera specific for the *osa* protein. As an alternative, we will also add a small epitope tag to the *osa* ORF and reintroduce the altered gene into *Drosophila* by P-element-mediated transformation. This epitope-tagged protein will provide an alternate method for the purification of *osa* protein from embryonic nuclear extracts if our attempts to generate an antisera directed against the *osa* protein are unsuccessful. Our collaborator, John Tamkun, has already generated epitope-tagged *brahma* strains and antisera directed against the *brahma* protein. Based on our genetic observations, our first experiments will be directed towards determining whether the *brahma* and *osa* proteins physically interact by co-immunoprecipitation from nuclear extracts and by localizing the protein products on polytene chromosomes in larval salivary glands. We will also use

the the yeast two-hybrid system to examine physical interactions between the cloned *brahma* and *osa* proteins.

We will also continue to characterize the other mutations that show strong genetic interactions with both *brahma* and *osa* mutations. In particular, we will focus on the molecular characterization of a gene called *Pearl*. Our preliminary mapping data show that *Pearl* is in the same chromosomal region that includes the *Drosophila* homologue of a *brahma*-related gene in yeast and mammals. Mutations in this *brahma*-related yeast gene appear to enhance mutations in the yeast *brahma* homologue.

Finally, we will continue our molecular and developmental analyses of the *l(3)87Ca* gene. We will determine the extent of the transcription unit, the sizes of the transcripts, and the spatial and temporal patterns of expression for the transcripts. We will also isolate cDNA clones corresponding to the transcripts and determine the putative protein sequence. The mutations will be used to determine in which cells and during which developmental stages *l(3)87Ca* gene expression is required.

Publications:

Felsenfeld AL, Kennison JA. Positional signaling by *hedgehog* in *Drosophila* imaginal disc development, *Development* 1995;121:1-10.

Hellmich MR, Kennison JA, Hampton LL, Battey JF. Cloning and Characterization of the *Drosophila melanogaster* CDK5 homolog, *FEBS Letters* 1994;356:317-321.

Kennison JA. The Polycomb and trithorax group genes of *Drosophila*: Trans-regulators of homeotic gene function, *Annual Reviews of Genetics* 1995;29:In press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01006-07 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Protein-Nucleic Acid Interactions in Vertebrate Embryogenesis

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

P.I.: T. Sargent Research Microbiologist (All listed LMG:NICHD)
 Others: R. Winning Visiting Fellow S. Campbell Biologist
 M. Morasso IRTA Fellow C. Nocente Prof. Serv. Contract
 K. Deonauth Pre-IRTA Student
 J. Franklin IRTA Fellow
 J. Scales NRC Fellow

COOPERATING UNITS (if any)

K. Mahon, LMGD, NICHD

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Unit on Vertebrate Development

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

6.08

PROFESSIONAL:

5.08

OTHER:

1.0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The general aim of this project is to identify and study the function of factors that regulate vertebrate embryonic development. This is being approached in three ways. One approach is to analyze the regulation and function of one of the vertebrate homologs of the *Drosophila* gene *Distal-less*. In flies this gene regulates limb and sensory appendage formation. The frog and mouse homologs are both expressed in epidermis, branchial arches, and in gradients in the limbs. The mouse gene is also active in hair follicles, tooth primordia and other locations. DNA elements and protein-DNA interactions responsible for this expression pattern are being mapped. In addition, targeted inactivation and mis-expression studies are being used to test for the function of *Distal-less* gene expression in the epidermis of transgenic mouse embryos.

A second approach employed has been to clone mRNAs encoding receptor-type protein kinases that are expressed in specific tissues of early frog embryos. One such kinase, called Pagliaccio, has been shown to affect cell-cell adhesion when constitutively activated in *Xenopus* embryonic ectoderm. A related gene, TCK, is also being studied.

The third approach focuses on the developmental function of calcium-dependent cell adhesion molecules in the zebrafish neural tube. A partial cDNA encoding a novel cadherin, VN-cad, has been cloned, sequenced, and characterized by in situ hybridization. When the complete protein coding region has been obtained for this gene, functional studies will be carried out in zebrafish embryos.

Project Description

Objectives:

The overall objective of this project is to elucidate molecular mechanisms that regulate vertebrate development. Specifically, this is approached by studying a small number of selected genes which, for different reasons, we believe have important roles in development. These are XDII-2, a *Xenopus* homolog of the *Drosophila* homeodomain gene *Distal-less*, *Pagliaccio* and *TCK*, which are members of the Eph class of receptor tyrosine kinases and are expressed in unique patterns in the early *Xenopus* embryo, and *Ventral Neural Cadherin* (VN-Cad), a calcium-dependent cell adhesion molecule expressed in the ventral neural tube of zebrafish embryos.

In addition to characterizing the expression pattern of these genes, their developmental functions are elucidated by introducing ectopic or constitutively active expression in transgenic embryos. Disruption of normal function is also achieved by targeted deletion of the endogenous gene and by dominant negative gene expression or antisense methods. The effects of these treatments on the affected embryos is assessed at the tissue, cell and biochemical levels.

Methods Employed:

An extensive array of molecular cloning techniques is used to generate and modify the genomic and cDNA clones used in this study. Micro-injection of DNA into cleaving blastomeres of *Xenopus*, and transgenic mice are the primary functional assays for regulatory elements. The function of specific protein molecules are tested by injecting synthetic mRNAs that have been engineered to disrupt the signal transduction pathways used by the products of the endogenous gene. The effects this has on development and on embryonic cell-cell communication is monitored by *In situ* hybridization, immunocytochemistry, and reverse-transcription PCR procedures. Developmental functions are also revealed by ectopic expression of regulatory genes such as *Distal-less* in inappropriate tissues, which can result in disrupted tissue differentiation. DNA-protein interactions are evaluated by gel mobility shift and footprinting assays. Functional aspects of transcription factors are tested by co-transfection of expression and reporter constructs into tissue culture cells, and by *in vitro* transcription. Classical embryological techniques such as microdissection, cell dissociation, and tissue recombination are also used. Protein-protein interactions are identified by two-hybrid screening in yeast and by co-immunoprecipitation assays. Tyrosine phosphorylation is detected by western blots with antiphosphotyrosine antibodies, usually following immunoprecipitation.

Major Findings:

Ectopic expression of *Distal-less* in basal keratinocytes severely disrupts skin development in transgenic mouse embryos.

Constitutive activation of the receptor tyrosine kinase *Pagliaccio* in *Xenopus* embryos results in dissociation of ectodermal cells.

The signal transduction processes associated with *Pagliaccio* activation and the resulting loss of cell adhesion do not require functional ras activity.

The expression of the ventral neural marker VN-cadherin is not grossly disrupted in homozygous mutant zebrafish embryos lacking floor plate (*cyclops*) or notochord (*no tail*). This supports the hypothesis that these structures are not a prerequisite for patterning in the neural tube.

Significance to Biomedical Research and the Program of the Institute:

Xenopus, mouse, and zebrafish embryos are versatile and complementary models for the study of vertebrate development. Each system has unique attributes that facilitate certain experimental approaches; for example, gene transfer into the mouse is an excellent method for testing the function of developmentally significant molecules, and the Xenopus embryo is ideally suited for studying cell-cell interactions during development. The zebrafish is highly amenable to broadly based mutagenic screens, and can also be physically manipulated in the same way as Xenopus. Together these three systems make a powerful combination that can be used to identify basic and phylogenetically conserved mechanisms controlling vertebrate, and thus human, embryogenesis and tissue differentiation. The studies carried out in this project should help elucidate molecular mechanisms regulating skin development and homeostasis, and nervous system pattern formation. Insights gained into the control of cell-cell adhesion should also be relevant to epithelial-mesenchymal interactions occurring during development and to the invasive behavior of metastatic neoplasia. These are all relevant to human health and are areas of major importance to the mission of NICHD.

Proposed course of project:

Distal-less

The analysis of the Xenopus XDII-2 regulatory region will continue. Mobility shift assays have detected formation of specific complexes between this DNA and nuclear extracts from epidermal cells, and DNA footprinting will be carried out to precisely map the binding sites on the XDII-2 gene.

The primary focus of effort relating to Distal-less will be determining the function of this gene in epidermal differentiation. Our earlier finding that Dlx-3 expression in skin is limited to suprabasal cells suggested that Dlx-3 may regulate important aspects of keratinocyte differentiation into stratified skin. We will test this by forcing Dlx-3 expression in basal cells using a construct driven by the basal keratinocyte-specific keratin K5 promoter, introduced into transgenic mice. Preliminary results indicate a severe and lethal disruption of skin development in these animals. The nature of this disturbance will be investigated by a combination of molecular and histochemical tools, including antibodies recognizing an array of epidermal markers such as keratins K6, K8, K10 and K14, involucrin, filaggrin and transglutaminase.

In order to obtain sufficient material for this analysis, we plan to prepare transgenic mouse strains carrying a modified construct which is interrupted by a transcriptional attenuator bracketed by Lox recombination sites. Crossing such mice with Cre-producing strains should result in high penetrance expression of the introduced Distal-less gene in the basal cells of most, if not all progeny.

In collaboration with K. Mahon (LMGD/NICHD) we will disrupt the endogenous Dlx-3 gene by site-specific recombination in ES cells, followed by incorporation into chimeric mice which will be bred to produce heterozygous carriers of the disrupted gene. Homozygous null embryos will be generated by crossing these carrier mice, and affected embryos analyzed at the gross, histological and biochemical levels, with particular attention to skin morphology and differentiation.

Receptor Tyrosine Kinases.

The project aimed at studying protein kinases in *Xenopus* development has been focused to two genes, both receptor tyrosine kinases (RTKs) of the EPH class. One, Pagliaccio (Pag) is activated at early gastrula and is transiently expressed in several specific sites of migratory cell behavior or epithelial-mesenchymal transition (pronephros). Homologs of Pag have been identified in mammalian species, including humans. The other RTK being studied is TCK (tail/cement gland kinase), which is present as a maternal RNA. TCK expression becomes localized during early organogenesis.

To investigate the function of Pag, we will continue to utilize chimeric molecules consisting of the extracellular ligand-binding domain of the epidermal growth factor receptor (EGFR) fused to the transmembrane and intracellular domains of Pag. RNA encoding this hybrid molecule, as well as control RNA encoding a point mutant with a disabled kinase function, will be injected into the blastomeres of early frog embryos. The hybrid kinase is activated by co-injection of synthetic mRNA encoding human TGF α , a ligand for EGFR. We have shown this results in dissociation of the *Xenopus* embryo. We will use co-immunoprecipitation and western blot procedures to identify proteins that interact with Pag to elicit this phenotype. In addition, a two-hybrid screen will be carried out using the Pag intracellular domain as "bait" to identify genes encoded by *Xenopus* embryonic mRNA that can physically interact with this portion of the protein.

EGFR-Pag constructs are being used to establish permanent transfected NIH3T3 cell lines. An inducible expression system, based on reversible repression via *E. coli* lac repressor protein has been employed to circumvent the tendency of Pag to autoactivate when present at high protein concentrations.

These lines will be used in studies of the Pag signal transduction pathways. If this expression approach is successful with 3T3 cells, it will be extended to epithelial cell lines, such as the epidermal carcinoma A431 line, or mammary epithelium-derived tumor cell lines, which should help to clarify the mechanism by which Pag activity disrupts cell-cell adhesion.

Experimental overexpression of TCK has revealed that this kinase, while related to Pag, does not drastically alter cell adhesion in frog embryos. Further analysis of TCK will focus on identification of its protein phosphorylation and interaction targets. This will be carried out in the *Xenopus* embryo system, as well as in NIH3T3 cells and A431 epidermal cells, both of which express the mammalian homolog of TCK, Hek2.

VN-Cadherin

Continued investigation of VN-cad function will require isolation of a longer length cDNA clone that includes the entire VN-cad protein coding sequence. Repeated library screens have not yielded the desired clones, so the missing cDNA will be obtained by reverse transcription/PCR. Once the complete VN-cad protein sequence is available, mis-expression studies will be carried out in zebrafish. Synthetic mRNA will be injected into cleavage stage cells. The resultant embryos will be examined for neurological and other defects. Injected embryos will also be used as donors for cells to be transplanted into the prospective neural keel region of non-injected recipients, in order to test the hypothesis that VN-cad expression is involved in pathfinding for nerve bundles.

Finally, mouse homologs of VN-cad have been cloned by low-stringency screening of an embryonic (11 dpc) cDNA library. These cDNAs will be partially sequenced and analyzed by northern blot and

in situ hybridization, to determine if a true homolog has been identified. If this is the case, future plans may include mis-expression and/or targeted disruption of the mouse gene.

Publications

Morasso MI, Mahon, KA, Sargent, TD. A *Xenopus* Distal-less gene in transgenic mice: Conserved regulation in distal limb epidermis and other sites of epithelial-mesenchymal interaction. *Proc Natl Acad Sci USA* 1995;92:3968-3972.

Scales JB, Winning RS, Renaud CS, Shea LJ, Sargent TD. Novel members of the Eph receptor tyrosine kinase subfamily expressed during *Xenopus* development. *Oncogene*, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01008-06 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Molecular Genetics of Protein-Nucleic Acid Interactions in *Drosophila*

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

P.I.:	Susan R. Haynes	Expert	LMG:NICHD
Others:	Virginia Heatwole	IRTA Fellow	LMG:NICHD
	Cecilia Pazman	Visiting Fellow	LMG:NICHD
	Stefan Pype	Visiting Fellow	LMG:NICHD
	David T. Stollow	NRC Research Associate	LMG:NICHD
	Monica Cooper	Biologist	LMG:NICHD
	Colin Steven	Summer Student	LMG:NICHD

COOPERATING UNITS (if any)

Ann Beyer, Department of Microbiology, University of Virginia Medical School

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Unit on Protein-Nucleic Acid Interactions

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

4.65

PROFESSIONAL:

3.55

OTHER:

1.1

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The goal of this project is to understand the roles of RNA binding proteins in the post-transcriptional regulation of gene expression, through an analysis of the in vivo functions of proteins that contain RNA recognition motif (RRM) domains. RRM domains have been identified in many proteins involved in post-transcriptional regulation, and can mediate sequence-specific binding to RNA targets. We use genetics, biochemistry and molecular biology to study six *Drosophila* RRM proteins. Two of these, TSR and RB97D, are expressed during spermatogenesis, and are required for normal male fertility. The *Tsr* gene is expressed only in the male germ line, and mutations in *Tsr* drastically reduce male fertility. Electron microscopic analysis of testes from *Tsr* mutants reveals dramatic defects during spermatid morphogenesis. The mitochondria develop abnormally, with an aberrant relationship to the axoneme of the sperm tail. TSR is a cytoplasmic protein and is present at high levels in older primary spermatocytes and young spermatids. Translational regulation is known to be critical during spermatogenesis, and TSR may be involved in this process. We have used the yeast two-hybrid system to identify other proteins that may interact with TSR. One protein that specifically interacts with TSR is the ADP-ribosylation factor (ARF), a small GTP binding protein. ARF proteins have been implicated in membrane trafficking and organelle integrity, and a role for ARF in the remodeling of the spermatid mitochondrion is plausible. The TSR-ARF interaction might serve to integrate mitochondrial morphogenesis with the post-transcriptional regulation of specific mRNAs. Mutations in the *Rb97D* gene cause male sterility. *Rb97D* encodes a nuclear protein that binds to the *ks-1* lampbrush loop of the Y chromosome in primary spermatocytes. The *ks-1* loop is a transcriptionally active region containing a gene essential for male fertility. The binding of RB97D is RNase-sensitive, suggesting association with *ks-1* transcripts. When transcription ceases and the loop regresses late in spermatogenesis, RB97D is no longer detectable. These data are consistent with the idea that RB97D is required for processing nascent transcripts from the *ks-1* gene.

Project DescriptionPersonnel:

S. Haynes	Expert	LMG, NICHD
M. Cooper	Biologist	LMG, NICHD
V. Heatwole	IRTA Fellow	LMG, NICHD
C. Pazman	Visiting Fellow	LMG, NICHD
S. Pye	Visiting Fellow	LMG, NICHD
C. Steven	Summer Student	LMG, NICHD
D. Stolor	NRC Fellow	LMG, NICHD

Objectives:

Despite major advances in recent years, there are still many unresolved questions regarding the mechanisms of gene expression and regulation, particularly regarding the involvement of post-transcriptional processes in gene regulation. Understanding these processes will require a detailed knowledge of the protein-nucleic acid interactions that are involved. We are interested in analyzing the *in vivo* roles of RNA binding proteins in post-transcriptional regulation. The fruit fly *Drosophila* is a useful organism for such studies, since it offers the opportunity to apply genetics as well as biochemistry and molecular biology to the study of these processes in a multicellular eukaryote.

Methods Employed:

Culture and genetic analysis of *Drosophila* are done by standard methods. Creation of transgenic fly strains follows established procedures. Genes are localized by *in situ* hybridization to polytene chromosomes and by hybridization of isolated DNA fragments to Southern blots of genomic DNA from deletion chromosomes. Analysis of gene expression within developing embryos follows protocols developed by Tautz for use of digoxigenin-labeled probes for whole mount *in situ* hybridizations. Standard recombinant DNA technology is employed to screen phage cDNA and genomic DNA libraries for clones of interest, and to isolate, subclone, restriction map and sequence the DNA. Demonstration of RNA binding activity is done by UV cross-linking and gel shift analysis. Identification of RNA targets of RNA binding proteins is done by *in vitro* selection assays (SELEX) using *in vitro* transcription and PCR amplification of oligonucleotides that bind with high affinity. Preparation and analysis of RNA from different stages of development follows standard procedures. Antibodies directed against fusion proteins or peptides are prepared by standard means. Proteins are analyzed on Western blots of one- and two-dimensional gels. Yeast two-hybrid systems are used to identify interacting proteins.

Major findings:

In the cell, RNA is always associated with proteins, which are involved in its packaging, processing, transport, stability and translation. Many of the proteins involved in these processes have a common, well-conserved RNA binding domain termed the RNA recognition motif (RRM). Our studies have focussed on six RRM proteins. Two of these proteins, *Hrb87F* and *Hrb98DE*, are components of *Drosophila* hnRNP complexes, which consist of newly synthesized premessenger RNA associated with a specific set of nuclear proteins. These complexes and their component proteins have been extensively characterized in vertebrates, but the functions of individual proteins are largely unknown. The *Drosophila* *Hrb* proteins have two copies of the RRM and a glycine-rich C-terminal domain. The *Rb97D* and *Tsr* proteins also have similar RRMs, but differ from the *Hrb* proteins and

Appendix

1. The first part of the appendix contains a list of the names of the persons who have been elected to the office of Mayor of the City of New York since the year 1784.

2. The second part of the appendix contains a list of the names of the persons who have been elected to the office of Mayor of the City of New York since the year 1784.

3. The third part of the appendix contains a list of the names of the persons who have been elected to the office of Mayor of the City of New York since the year 1784.

4. The fourth part of the appendix contains a list of the names of the persons who have been elected to the office of Mayor of the City of New York since the year 1784.

5. The fifth part of the appendix contains a list of the names of the persons who have been elected to the office of Mayor of the City of New York since the year 1784.

each other in their C-terminal domains. Both proteins are required for normal male fertility. Finally, the *caz* and *Nts* proteins have single RRM domains which differ significantly from those of the other four proteins. *caz* is the *Drosophila* homolog of the human EWS and TLS genes. In certain sarcomas, translocations cause an in-frame fusion of EWS or TLS to the DNA binding domains of other proteins. We have initiated a screen for mutations in *caz* to assess its role in flies. The function of *Nts* is also being assessed, using a strain with a deletion near the 5' end of the gene.

TSR. We have previously reported on our initial genetic and molecular characterization of the *Testis-specific RRM protein (Tsr)* gene. These results are briefly summarized here. The *Tsr* gene is transcribed only in males, beginning in the third larval instar. It encodes a male germ line-specific RNA binding protein involved in spermatogenesis. In the absence of this protein, most sperm fail to complete development, and instead degenerate in the testis. The TSR protein is 46.6 kDa and has two RRM domains that are most closely related to the RRM domains of the *Hrb* proteins. The C-terminal half of the protein is unrelated to the glycine-rich domain of the *Hrb* proteins, or to any other protein in the sequence database.

TSR is required for normal male fertility, although males that are homozygous for a null mutation may produce a few progeny. Most of the spermatids in the testes of mutant males degenerate before the completion of spermatogenesis. Some are found in the seminal vesicles, but most of these are nonmotile and nonfunctional. Staining of fixed preparations of testes with nucleic acid stains such as orcein shows aberrations in young spermatids. Normally, spermatids develop in bundles with the nuclei clustered at one end. In TSR⁻ males, nuclei are often scattered throughout the bundle, as though they have become detached from the growing axoneme. We have also examined thin sections of mutant and wild type testes in the electron microscope. The sperm tails are disorganized in the mutants, and the mitochondria develop abnormally and do not associate properly with the axoneme.

Using the pET expression system, we purified TSR protein from *E. coli*, and used it to produce anti-TSR antibodies. Whole-mount staining of *Drosophila* testes with the affinity purified antibody demonstrated that TSR is a cytoplasmic protein whose expression is restricted to a limited period of spermatogenesis. The protein is first detectable at low levels in young primary spermatocytes, and continues to accumulate in older primary spermatocytes. The highest protein levels are seen in cells immediately before and after meiosis. The protein is present throughout the cytoplasm of early elongating spermatids but disappears before elongation is complete.

We have used the yeast two-hybrid system to identify proteins that may interact with TSR. The most promising candidate to date, which has been isolated multiple times in several screens, is the ADP-ribosylation factor, ARF. ARF is a small GTP binding protein related to the RAS superfamily of GTP binding proteins. ARF proteins have been implicated in many important cellular processes, including signal transduction, membrane trafficking and organelle integrity. We have used deletion mapping assays to identify the interaction domains in each protein. Interestingly, the portion of TSR that interacts with ARF is the first RRM domain. Whether this interaction is compatible or competitive with RNA binding is not known. We detect strong interaction between these proteins in two different yeast two-hybrid systems, and can coimmunoprecipitate them from *E. coli* extracts.

Post-transcriptional regulation is very important during spermatogenesis, since most transcripts required after meiosis are synthesized premeiotically and stored as RNP complexes. Based on its expression pattern, TSR could be a component of such complexes. The identification of ARF as a TSR interacting protein is unexpected, but is particularly interesting in light of the proposed functions of ARF and the alterations in spermatid

mitochondrial morphogenesis that are seen in TSR⁻ mutants. We speculate that the TSR-ARF interaction might serve to integrate mitochondrial morphogenesis with the post-transcriptional regulation of specific mRNAs.

RB97D. The *Rb97D* gene encodes an RRM-containing protein that is expressed in the testis and is required for spermatogenesis. Although it is expressed in other tissues as well, a complete deletion of the gene has no obvious phenotype other than male sterility. Using the pET expression system, we produced a protein corresponding to the RRM domains of RB97D in *E. coli*. The purified protein was used to generate anti-RB97D antibodies. Using affinity purified antibodies, we have shown that RB97D is a nuclear protein that is expressed only in pre-meiotic primary spermatocytes. During spermatocyte development, three regions of the Y chromosome become transcriptionally active and cytologically visible as large lampbrush loops. These regions are essential for fertility, as deletions or mutations cause male sterility. The RB97D protein binds to the *ks-1* lampbrush loop in an RNase-sensitive manner, suggesting that it is associated with *ks-1* transcripts. In addition, the *ks-1* loop appears to form normally in the absence of RB97D, indicating that RB97D is not required for its transcription. The Y chromosome is completely heterochromatic, and genes on the Y are likely to consist of short exons interspersed with large regions of satellite sequences and transposable elements. RB97D could be associated with exon sequences corresponding to the *ks-1* gene, or with repetitive sequences in the introns. The former is perhaps more likely, since the anti-RB97D antibody stains one of the loops in the Y chromosome of the related species, *Drosophila hydei*, and the Y chromosomes of the two species are known to contain different repetitive elements. Although other proteins that bind to Y chromosome loops have been identified (most of them only by immunological techniques), RB97D is the first whose expression is known to be required for spermatogenesis. The *Rb97D* mutant phenotype is similar to that of deletions of the *ks-1* gene. Its expression pattern and the fact that it is an RRM protein are consistent with the idea that it is required for processing nascent transcripts from the *ks-1* gene.

NTS. We have continued our molecular analysis of the *Nts* gene, which had been initiated last year. *Nts* is located in a cluster of RRM proteins, which includes *Hrb87F* and *Tsr*, on the right arm of the third chromosome at 87F. Northern blots show that *Nts* is transcribed throughout development, with the highest levels present in early embryos. At least four transcripts are detectable in early embryos. We have isolated multiple cDNA clones, and shown that the differences between the transcripts are likely to be due to use of alternative exons and polyadenylation sites. So far, no alternative processing that affects the coding region has been detected. *Nts* encodes a 75 kDa protein with a single RRM located near the C-terminus. This RRM shows limited sequence homology to the RRMs of the other proteins in the 87F region; its sequence suggests that it belongs to a distinct subfamily of RRM proteins. Much of the N-terminal region of NTS consists of an extensive glutamine-rich domain. Preliminary genetic experiments with a fly strain having a deletion in the putative promoter region of *Nts* suggest that it is required for both male and female fertility. We have expressed portions of the NTS protein in *E. coli*, and are using the purified protein to generate antibodies for immunolocalization studies.

Significance to Biomedical Research and the Program of the Institute:

A major aim of the Institute research program is an understanding of the molecular mechanisms of the regulation of gene expression. RNA binding proteins play critical roles in the metabolism of RNA transcripts. *Drosophila*, because of the applicability of genetic, molecular and developmental techniques in this organism, offers significant advantages in the study of such problems. The identification of two RRM-containing proteins that regulate alternative splicing in the *Drosophila* sex determination pathway, and biochemical evidence for the involvement of mammalian hnRNP proteins in splicing, have stimulated interest in the study of such proteins as important regulatory molecules.

Proposed course:

Tsr. A major effort will be directed toward assessing the biological relevance of the TSR-ARF interaction. We have obtained a *Drosophila* strain that may have a P element integrated near the *Arf* gene. This will be used to generate *Arf* mutants, by imprecise excision or local hopping of the P element. The mutants will be used to assess the role(s) of ARF in spermatogenesis, and to look for interactions with *Tsr*. We will also use chromosomal duplications to vary the dosage of *Arf* relative to *Tsr* and examine spermatogenesis under these conditions. We have generated a stable cell line expressing TSR under the control of an inducible promoter, which will be used for further biochemical characterization of the TSR-ARF interaction. We will use sucrose gradient sedimentation of testis extracts to determine whether TSR is present in polysomal mRNP complexes, in free mRNP, or is not complexed with other proteins or RNAs in vivo.

Rb97D. We will continue our analysis of the binding of RB97D to the *ks-1* Y chromosome loop. A major goal will be the identification of the RNA target(s) of RB97D. The gene encoded by *ks-1* is unknown. If RB97D is required for processing *ks-1* transcripts, these transcripts might be unstable or structurally altered in the absence of RB97D protein. We will use RNA from wild type and RB97D⁻ mutant males in differential display and representational difference analyses to detect transcripts that are altered in the mutants. These can be rapidly checked to determine whether they are transcribed from *ks-1*. Candidate transcripts will be tested for specific high-affinity binding to RB97D, and for alterations in RB97D⁻ mutant males.

Nts. We are currently producing anti-NTS antibodies, and will use them to determine the tissue distribution and subcellular localization of the NTS protein. We will characterize the existing deletion mutation near *Nts* to determine its effects on *Nts* expression, and study the phenotypic effects of this mutation. Preliminary results suggest that mutations in *Nts* are not lethal, but affect male and female fertility. If the results of our initial experiments are encouraging, we will begin a genetic analysis to obtain additional mutations in *Nts*.

cabeza. We have begun to use the yeast-two hybrid system to identify interacting proteins, and positive clones from these screens will be studied. We also have generated the appropriate strains to use in a P-element screen for insertions into the *caz* gene, to begin a genetic analysis of its function.

Hrb87F and Hrb98DE. Future work on the *Hrbs* will be directed toward understanding their interactions with each other. This interaction appears to be important not only in spermatogenesis but also in oogenesis. We have a possible mutation in *Hrb98DE* that was generated in a screen for interacting genes. We will use genetic and molecular tests to determine if it truly is a *Hrb98DE* mutation. If so, we will analyze its phenotype alone and in combination with *Hrb87F* mutations.

Publications:

Haynes SR. Posttranscriptional regulation and RNA binding proteins in development. J Biomed Sci 1995; in press.

Stolow DT, Haynes SR. Cabeza. A *Drosophila* gene encoding a novel RNA binding protein, shares homology with EWS and TLS, two genes involved in human sarcoma formation. Nucleic Acids Res 1995;23:835-43.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01009-03 LMG

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Regulation and Function of Genetic Elements

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	H.L. Levin	Senior Staff Fellow	LMG:NICHD
	A. Atwood	Bio Lab Tech	LMG:NICHD
	J. Lin	IRTA	LMG:NICHD
	M. Williams	Visiting Fellow	LMG:NICHD
	J. Choi	Pre-IRTA	LMG:NICHD

COOPERATING UNITS (if any)

Dr. Jef Boeke, Johns Hopkins University, Baltimore, Maryland
 Dr. Michael Benedik, University of Houston, TX

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Unit on Yeast Molecular Genetics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

4.3

PROFESSIONAL:

4.0

OTHER:

0.3

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Our research program focuses on the mechanisms of retroelement action. Our approach to understanding the complex interactions between the retroelement and its host is to study retrotransposons, a family of elements that are closely related to retroviruses. A significant advantage to studying retrotransposons is they exist in hosts such as yeast that can readily be studied using sophisticated molecular genetic techniques. In the process of characterizing yeast transposition, we have collected strong evidence that Tf1 reverse transcriptase uses a novel self-priming mechanism to initiate cDNA synthesis. This is in complete contrast to the tRNA mechanisms thought to be used by all other LTR-containing elements. In this report we describe the characterization of the minus-strand strong-stop DNA that provides additional support for the self-priming mechanism. Genetic and biochemical analysis of Tf1 RT mutations in the active site of the polymerase allowed us to observe priming intermediates consisting of transcripts that had the first 11 bases removed. This data suggested a molecular model for priming that includes a cleavage of the first 11 bases of the transcript and the priming of reverse transcription from the 3'OH of the 11th base.

The analysis of a large family of point mutations near the primer binding site (PBS) confirmed the presence of a new 39 base pair RNA structure that is essential for transposition. As this large structure includes the 11 base pairs shown to be important for priming, we speculate that the newly detected structure may also participate in the self-priming mechanism. The assembly of functional Tf1 particles has been a paradox since other retroelement particles assemble with a molar excess of capsid protein that accumulates because the levels of the Pol proteins are restricted by reading frameshifts or stop codons. Tf1 however, expresses all its protein from within a single open reading frame as a primary translation product and we found that Tf1 particles contain a 26-fold excess of Gag compared to IN protein. By looking at cultures in different stages of growth, we have been able to observe an IN degradation process that leads to this excess of Gag. In addition, we found that most of the IN degradation occurred before Tf1 cDNA is synthesized indicating that the particles with a 26-fold excess of Gag are intermediates in transposition. We have used immunoblot analysis to reveal that IN degradation occurred in cells starved for glucose and not in those cell starved for nitrogen, suggesting that the loss of IN is not a time-dependent process but occurred only after certain growth conditions are met. To determine the factors required for transposition, we have created a large set of mutant strains that are defective for transposition. Thus far, we have identified six host genes that are required for either protein accumulation, particle stability, or integration.

Project Description

Objectives:

Our efforts are designed to elucidate the molecular basis of each step in the process of retroelement proliferation. The most medically important class of these elements are the retroviruses, members of which are responsible for AIDS as well as several types of cancer caused by the proviral induction of oncogene expression. However, many aspects of particle assembly, reverse transcription, nuclear entry and chromosome integration are difficult to study because retrovirus hosts possess a high level of genetic complexity. Our approach to understanding these events is to study retrotransposons, a family of elements that are closely related to retroviruses. A significant advantage of studying retrotransposons is they exist in yeast, a host that can be studied readily by using sophisticated molecular genetic techniques. The similarities of retrotransposons to retroviruses includes the presence of two long terminal repeats (LTRs) and open reading frames (ORFs) with coding sequences homologous to retroviral protease (PR), reverse transcriptase (RT), and integrase (IN). The first step in the transposition pathway is synthesis of a full-length mRNA with sequence that begins in the 5' LTR and terminates in the 3' LTR. This is directly analogous to the initial step in retrovirus particle formation. Retroviral and retrotransposon mRNAs are translated into proteins that assemble along with the mRNA into large particle structures. Both retroviral and retrotransposon particles undergo a maturation process that includes the proteolytic processing of precursor proteins and the reverse transcription of the mRNA. Retroviral particles are able to escape the host cell and integrate their DNA into the genome of neighboring cells, while retrotransposon particles complete transposition by simply inserting their DNA into the genome of the original host cell. Because each step in retrotransposition is directly related to a retrovirus process, results from the investigation of yeast retrotransposition will be relevant to aspects of retrovirus behavior.

The retrotransposon we study is the Tfl element of the fission yeast, *Schizosaccharomyces pombe*. The transposon is 5 kb and contains a single ORF of 1340 amino acids with coding sequences homologous to PR, RT, and IN. We have previously demonstrated, using an *in vivo* assay, that at least one of our cloned copies of Tfl is active and can transpose at a significant frequency. The system we have developed allows us to use a battery of powerful molecular and genetic techniques to identify and characterize factors that contribute to the transposition process. We are initially focused on the interactions between the host cell proteins and the element because it is these that are the least understood in the retrovirus systems. Genetic experiments on retroviruses have defined the principal role of each of the major virus proteins by altering the virus encoded factors and assaying for an effect on virus functions. Studying retroelements in yeast allows us to explore the contributions of the host proteins to the entire process of transposition.

The genetic approach we are using to identify proteins that function in transposition is to use a new form of the transposition assay to screen individual colonies for mutations that alter the frequency. We developed a genetic screen at NIH for transposition that depends on the overexpression of Tfl transcripts containing a selectable marker that is reverse transcribed during transposition and inserted into the host genome. We experimented with various combinations of growth media and expression systems with the result that we increased our previously reported transposition frequency of 0.5% per 10 cell generations to 18%. This improvement allowed us to develop this new transposition assay that is now being used to screen large numbers of cells for mutations that alter the transposition frequency. Mutations in genes that contribute to transposition will be used to clone the genes required for transposition from an *S. pombe* genomic library.

Another goal associated with the mechanisms of transposition is to learn which parts of the Tfl proteins are required for mobility and what are their roles in the process. Towards this goal, we have started to assemble a large set of random mutations in Tfl itself that block transposition. By characterizing the positions and properties of these mutations, we hope to associate each protein with a part of the transposition pathway and perhaps attribute new functions to some of these domains.

Further characterization of Tfl protein expression is also a goal because of its unique aspects among LTR-containing retroelements. Retroviruses and LTR-containing retrotransposons encode their Gag (capsid) and Pol (PR, RT, and IN) proteins in separate ORFs that are interrupted by a frameshift or a stop codon. This organization causes a molar excess of Gag to Pol protein to be expressed which is important for the assembly of particles that normally contain as much as 20-200 times more Gag than Pol. Tfl coding sequence for Gag and Pol is contained within a single ORF and we have used three separate experiments to show that all Tfl protein is expressed within a single primary translation product. Last year we determined that the Tfl proteins do form virus-like particles similar to those made by the other retrotransposons studied to date. One of our present objectives is to determine the ratio of Gag to Pol proteins within these particles and develop a model for how these ratios are achieved as expressed from a single ORF.

A unique aspect of Tfl is that its sequence contains no homologies to tRNA molecules at the site where reverse transcription is initiated, the primer binding site (PBS). This is in direct contrast to all other LTR-retroelements which have 15-20 bases of perfect homology to specific tRNA molecules that serve as primers for reverse transcriptase during the first step of cDNA synthesis. Although we have detected the presence of an RNA primer at the Tfl PBS, it appears not to be a tRNA. Our goal is to determine the identity of the reverse transcriptase primer as well as the mechanism used. Recent results indicate that Tfl uses a novel selfpriming mechanism of reverse transcription that defines a new family of retroelements.

Methods Employed

We have used a wide variety of molecular, genetic, and biochemical techniques in experiments described in this report. Our genetic approaches include several standard yeast techniques such as transformation, gene replacement, mating, tetrad dissection, diploid formation and mutagenesis. We are also using the two-hybrid system as a general method for identifying any *S. pombe* proteins that interact with Tfl proteins.

Our standard transposition assays take advantage of two drug marker selections. We use 5-fluoroorotic acid to select for cells that lose plasmid copies of the URA3 gene. Our transposons are marked with the bacterial *neo* gene that provides *S. pombe* with resistance to high concentrations of G418. We have also started to use an artificial *S. pombe* intron placed within marker genes in the transposon to assay for reverse transcription events.

Biochemical methods include the use of various yeast expression systems to improve production of the Tfl proteins. We now use the *nmt1* promoter of *S. pombe* to regulate high levels of protein synthesis. The Tfl ORF is fused to the *nmt1* promoter in a multicopy plasmid. The presence of 10 μ M vitamin B1 reduces expression 50-fold. *nmt1* is the strongest *pombe* promoter thus far characterized. Extracts from these expression strains are then subjected to sucrose gradient sedimentation as well as other purification protocols. Fractions from these gradients are immunoprecipitated by using antibodies raised against Tfl proteins.

A full range of molecular biology techniques are utilized to make reagents and characterize them. Constructs are made by using restriction enzymes, phosphatases, DNA polymerases, PCR and ligase. We routinely sequence the DNA of complex constructs. Immunoblots, DNA blots and RNA blots are all used to characterize various strains.

Major Findings:

A. Priming of Tfl reverse transcription.

1. Evidence that Tfl uses an unusual mechanism of self-primed reverse transcription.

The role of tRNA in the priming of reverse transcription is thought to be common to all LTR-containing elements since the retroviruses and LTR-retrotransposons that have been sequenced have a PBS that is complementary to known tRNA species (Figure 1A). In light of the conservation of reverse transcription priming, it is surprising that no tRNA primer has been identified for Tfl. In addition to the absence of tRNA homology, Tfl also lacks the highly conserved UGG in the PBS of retroelements that hybridizes to the last 3 bases of the tRNA molecule. Evidence for the presence of an RNA primer that functions precisely at the conventional PBS location of Tfl came from previous analyses of DNA isolated from virus-like particles. This DNA was used as template in primer extension analyses and treatment with RNase demonstrated that there are at least 9 bases of RNA at the 5' end of the minus strand strong-stop DNA. The position of this RNA is immediately downstream of the 5' LTR and defines the PBS for Tfl. Instead of postulating that an uncharacterized tRNA serves as the primer, we proposed in our last annual report that Tfl mRNA undergoes a novel self-priming mechanism of reverse transcription. The first 11 bases of the Tfl transcript are exactly complementary to the PBS location identified by the primer extension experiments. This observation raised the possibility that the first 11 bases of the transcript folds back and anneals to the PBS. An RNA endonuclease would be able to cut the first 11 bases from the transcript so that the 3' OH required to prime DNA synthesis would become available (Figure 1B). Alternatively, RT might prime DNA synthesis from a 2'OH of an internal nucleotide as has been reported for the bacterial retron reverse transcriptase [Hsu, 1992 #1424; Lampson, 1989 #836].

Experiments described in our last annual report showed that the mutations in Figure 1C at the PBS or the 5' end of the transcript reduced transposition more than 30-fold. When two mutations were combined to reestablish complementarity, transposition frequencies were restored to 80% of the wild type level [Levin, 1995 #1575]. Consistent with the hypothesis that the mutations block priming, we found that the individual mutations resulted in a drop in the levels of minus-strand strong-stop DNA as isolated from Tfl particles. Strains with the compensating double mutations produced wild type levels of strong-stop DNA.

2. Analysis of minus-strand strong-stop DNA.

In all cases where the strong-stop species was observed on DNA blots, the signal was a doublet composed of two closely migrating bands. In hopes of learning more about the priming mechanism, the doublet nature of the minus-strand strong-stop species was investigated. We used DNA blot analysis of sequence gels in combination with radiolabeled probes consisting of several oligonucleotides to map the 5' and 3' termini of the two strong-stop species. The sequence ladders on these gels indicated that the size of the two strong-stop species differed by about 11 bases (± 1 base). The results indicated that the larger band is the full-sized strong stop species with a 3' end that corresponded with the first

base of the Tfl transcript and a 5' end that has the RNA primer with the PBS sequence. The smaller species has the same RNA primer at the 5' end but the 3' end is missing the 11 bases that are templated by the beginning of the Tfl transcript. The shortened strong-stop molecule is likely due to a premature termination of reverse transcription resulting from either a cleavage of the mRNA or an attachment of a 2'-5' linkage at the 11th base of the Tfl transcript. Regardless of which of the two events caused the termination of reverse transcription, the presence of the shortened strong-stop product provides support for the self-priming model.

3. *In vivo* experiments to distinguish between two types of self-priming mechanisms Despite the genetic and biochemical evidence that the folded RNA structure is important for priming, several questions about the mechanism are left unanswered. One very significant question is how does the RNA structure allow for priming given that no 3' OH is made available by the formation of the fold-back loop. A likely possibility is that an endonuclease is responsible for cleaving off the 1st 11 bases of the Tfl transcript so that the 3'OH of the 11th base can serve as the site of priming. Another explanation for the lack of a 3' OH is that the Tfl reverse transcriptase may actually prime DNA synthesis from a 2' OH forming a branched intermediate similar to that produced by the bacterial retron reverse transcriptases [Hsu, 1992 #1424; Hsu, 1990 #998]. Our observation that minus strand strong-stop DNA synthesis is blocked near the 11th base of the Tfl transcript supports the idea that either a cleavage event or a 2' OH linkage occurs at the 11th base of the transcript.

To explore the nature of the priming mechanism, we mapped the 5' ends of Tfl transcripts extracted from particles to test for the presence of Tfl mRNA missing its first 11 bases. The result was that large amounts of Tfl transcript were detected, but all of it had 5' ends at the previously mapped 5' start site [Levin, 1992 #1196]. One complication of this result is that the putative cleavage event is likely coupled to reverse transcription so that if transcripts are cleaved, they would immediately be degraded by RNaseH. In addition, our inability to detect evidence for Tfl transcripts connected to strong stop DNA by a branch point might have been due to RNaseH degradation of the mRNA.

To preserve priming intermediates from degradation, we used site-directed mutagenesis to generate five Tfl plasmids each with a mutation in an active site residue of RNaseH. If these mutations greatly reduce RNaseH activity, any cleaved transcripts would be protected from degradation and thus could be observed by S1 nuclease mapping or primer extension. In addition, mutations that leave RT polymerase activity intact would also allow us to observe Tfl transcripts attached by a branch point to strong-stop DNA if a 2' OH group is used for priming. Several groups have identified 3 amino acid positions within RNaseH of HIV and moloney leukemia virus that can be mutated to produce an RT with normal polymerase activity and no RNaseH function [Repaske, 1989 #1479; Mizrahi, 1994 #1478]. Using these studies as a guide, as well as the identification of active site residues from X-ray crystal structures of other RNaseH molecules [Davies, 1991 #1598; Katayanagi, 1990 #1599; Yang, 1990 #1600; Kohlstaedt, 1992 #1601], we individually changed two essential aspartic acids (443 and 498, HIV numbers) to asparagines and alanines while a glutamic acid (478) was changed to a glutamine. These mutations were incorporated into the Tfl transposition plasmid and their proteins expressed in our *S. pombe* strains. Transposition assays of the mutant Tfl elements were used to verify a lack of mobility, and *in vitro* RT assays of the Tfl particles were used to identify the mutations that preserve DNA polymerase activity. All 5 mutations caused a severe reduction in transposition frequency. Replacement of either of the aspartic acids with glutamine resulted in particles with wild type levels of RT activity as measured by *in vitro* assays that include synthetic primer and template. We found that the alanine replacements of the aspartic acids retained approximately half of the RT activity and the glutamine replacement of the glutamic acid had no detectable activity. Extracts from partially purified

particles were used to isolate Tfl mRNA to be examined for cleavage at the 5' end as well as for covalent linkage to strong-stop DNA. None of the RNA isolated from any of the mutant or wild type strains exhibited evidence of a 5' end cleavage as indicated by primer extension studies. These results ostensibly suggested that cleavage of the first 11 bases of the transcript was not required for self-priming. However, examination of the particles for DNA indicated that none of the mutant RT polymerases were able to produce minus-strand strong-stop material. Therefore, the lack of any observable cleaved mRNA could have been due to a complete lack of priming events caused by the RNaseH mutations. In addition, the absence of minus-strand strong-stop DNA in the mutants suggested that the priming event may actually require RNaseH activity.

An additional *in vivo* experiment designed to detect the generation of a primer via a cleavage event of the Tfl transcript avoided mutations in RNaseH in case it itself possesses the putative cleavage activity. We used versions of Tfl RT that have defective polymerase activity with the objective of trapping priming intermediates that have undergone the putative cleavage event but not extension. We expected that mutations in the polymerase domain of RT would allow the accumulation of cleaved Tfl transcripts because RNaseH would be unable to degrade the cleaved mRNA due to the lack of DNA:RNA duplex. 3 Tfl assay plasmids with mutations in the conserved YXDD box of the RT polymerase were created. RNA from mutant and wild type particles was examined by primer extension and the results indicated that 20%-50% of the RNA from the mutant particles had suffered a cleavage event that removed just the first 11 bases. Wild type particles showed no cleavages indicating that RNaseH was degrading the RNA as cleavage events occurred. S1 nuclease mapping experiments are in progress to independently examine the 5' end of the mutant RNA for cleavage sites. Should these confirm the presence of cleaved RNA in particles, we will combine the RT polymerase mutations with the RNaseH mutations to test the role of RNaseH activity in the cleavage event. A role of RNaseH in the cleavage of the Tfl self-priming RNA would be particularly interesting given recent reports that RNaseH from retroviruses possess an activity designated RNaseH* that results in cleavage of RNA in an RNA:RNA duplex [Hostomsky, 1994 #1385].

4. The 11 base pair RNA duplex required for self-priming comprises one third of a much larger structure of continuous RNA base pairs.

An important question left unanswered by our current self-priming model resulted from the close inspection of the RNA sequence in this leader region. Figure 2 shows a large potential RNA structure containing the 11 base pairs of the PBS as well as 28 other continuous base pairs. We are interested in defining the function of this large potential structure in priming as well as other processes such as packaging of Tfl mRNA into the virus like particles. To this end, we have randomly mutagenized the entire region of secondary structure by PCR and isolated a total of 78 mutations, 40 of which represent single base mutations that reduce transposition significantly. These strains have been subjected to our quantitative transposition assay and the results indicated that approximately half exhibited at least 50-fold less transposition activity while the others showed 5 to 50-fold lower transposition. Mutations of both classes were distributed throughout the PBS region, the adjacent 11 base pair region, and the 5 base pair section just upstream of the PBS. To address the effect of the mutations on expression, each of these Tfl elements was subjected to immunoblot analysis and found to synthesize wild type levels of Gag and IN protein.

We have recently developed a DNA blot method for the detection of Tfl cDNA that enables us to measure amounts of mature double-stranded reverse transcript. *S. pombe* cell extracts are phenol extracted and ethanol precipitated. The resulting nucleic acid is restriction digested with enzymes that

produce cDNA fragments with very different sizes from the plasmid Tfl sequences. This method was used to screen each of 40 different Tfl elements with mutations in the RNA leader structure. The results showed that each of the mutants produced lower levels of cDNA than wild type. When the amounts of cDNA were quantitated, the reduction in reverse transcription caused by each mutation correlated well with the magnitude of their transposition defect. Although this data suggests each mutation only effects priming levels, we are also testing the mutations for defects in the packaging of RNA into particles, another activity that would result in lower levels of cDNA synthesis.

B. The ratio of Gag compared to IN in Tfl particles.

1. The Tfl Gag protein accumulates to significantly higher levels than IN.

We have previously shown that the Tfl primary translation product is a single 140 kDa polypeptide that is cleaved by PR to form the mature Gag, PR, RT and IN proteins. Further characterization of the Tfl proteins is important because their expression is unique among LTR containing retroelements. Retroviruses and LTR-containing retrotransposons encode their Gag (capsid) and Pol (PR, RT, and IN) proteins in different ORFs that are separated by a frameshift or a stop codon. This organization causes a molar excess of Gag with respect to Pol protein to be expressed because a small number of ribosomes are able to translate past the stop codon of Gag. Excess Gag is important for assembly of particles because Gag serves as a structural component of the particle. Since the Tfl coding sequence for Gag and Pol is contained within a single ORF, there is no obvious mechanism for overproducing Gag protein. The high levels of transposition exhibited by Tfl and the particle nature of its proteins compelled us to develop methods for determining the ratio of Gag to Pol proteins in whole yeast extracts as well as in partially purified virus-like particles.

As reported last year, two different methods were developed to calibrate the immunoblot signals produced by anti-Gag and anti-IN antisera so that quantitative blotting could be used to evaluate the relative levels of these two proteins. Extracts from a PR⁻ *S. pombe* strain that contained only the 140 kD primary translation product were used to calibrate anti-Gag and anti-IN antisera since the Tfl product contained equal amounts of Gag and IN epitopes. Our second method for antisera calibration was based on Gag and IN protein expressed in bacteria and purified using Ni columns. Both methods for comparing the strength of the antisera generated similar results that indicated the anti-IN antiserum produced 1.5 to 3 times more signal than the anti-Gag antiserum. Using the calibrated antisera, we found immunoblots of *S. pombe* proteins showed that whole cell extracts as well as partially purified particles contained 26 times more Gag than IN. This significant excess of Gag in the Tfl particles represented a composition that resembled that found in retrovirus and retrotransposon particles [Stromberg, 1974 #1451; Panet, 1975 #1450; Farabaugh, 1993 #1389].

Despite the observation that Tfl particles appeared to assemble with the high levels of Gag expected for typical retroelement particles, the mechanism for adjusting the ratios of Gag to IN was very different than the frameshifting mechanism. We surmised that since Tfl proteins are all expressed at the same levels, a selective degradation mechanism must result in the reduction of IN levels. In support of this reasoning, we found that extracts made from log-phase yeast cultures contained equal amounts of Gag and IN while cultures of higher cell density contained the lower amounts of IN relative to Gag that are usually observed.

2. The analysis of Tfl particle development in *S. pombe* cultures.

This year we initiated detailed time-course experiments to carefully monitor the changes in IN levels as cells reached stationary phase. In agreement with observations reported last year, log phase cells in these cultures possessed approximately 1:1 ratios of Gag to IN and the stationary phase cells showed much less IN than Gag. Most of the IN degradation occurred within a 6 hour period between the OD₆₀₀ levels of 9.2 and 14.0. The final Gag to IN ratio observed in this experiment was 40:1. To determine if the Tf1 particles that contained excess Gag played a functional role in transposition, we examined the same cells from the time course for reverse transcription products. A blot of DNA extracted from whole cells, cut with *Bst*XI, and hybridized with a *neo* probe was used to measure the synthesis of Tf1-*neo* cDNA. The restriction cut allowed us to distinguish between *neo* sequences derived from reverse transcripts versus plasmids because Tf1 has a single *Bst*XI site 2 kb from the 3' end of the element. Additional evidence that the 2 kb band was derived from the Tf1 cDNA came from examination of an isogenic strain with a Tf1 plasmid that had a frameshift mutation in PR and therefore produced no RT. This mutant strain contained the plasmid band but not the 2 kb cDNA band. The 2 kb cDNA band derived from the double-stranded Tf1 cDNA was observed to increase greatly in the cells that contained reduced IN protein. The levels of the 2 kb band were quantitated and normalized to the amount of plasmid derived signal in each lane. The increase in Tf1 cDNA content observed during the entire time course was 8-fold and the amounts were linearly increasing even in cells from the last time point. The bulk of this increase occurred after the Gag to IN ratios increased to the levels observed in stationary phase. The fact that most of the cDNA is produced following degradation of the majority of IN protein was consistent with the hypothesis that particles containing a large molar excess of Gag compared to IN are functional in producing double-stranded cDNA. We cannot however determine whether the small amounts of cDNA produced in log phase cells are synthesized by particles of this type or by particles containing equal molar amounts of Gag and IN.

3. Examination of stationary phase cultures for conditions that trigger IN degradation.

The degradation of Tf1 IN in cells from stationary phase cultures is a regulated process that results in the high ratios of Gag to IN typical of retroelements. Several aspects of stationary phase conditions are candidates for initiating the degradation of IN. Nitrogen, sulfate, and glucose are all nutrients that may be limiting in stationary phase cultures and thus initiate IN degradation. The low pH of these cultures or the accumulation of ethanol are also properties of stationary phase that might result in IN degradation. To determine which if any of these conditions might effect the ratio of Gag to IN, we grew cells expressing Tf1 to midlog densities and placed them in media that contained different levels of the agents being tested. The results from this survey indicated that cells starved for nitrogen or sulfate exhibited the normal degradation of IN seen in stationary phase cells. However, we found that cells starved for glucose, were unable to degrade IN and therefore resulted in stationary cells that contained equal amounts of Gag and IN long after cells in complete medium exhibited the drop in IN levels. These results suggested that the turnover of IN may be a process that only occurs in cells arrested in G₀ since nitrogen and sulfate starvation of *S. pombe* causes G₀ arrest while glucose starvation causes a nonspecific arrest [Moreno, 1991 #1429]. Never the less, we plan to use plates with reduced glucose to assay the transposition frequency of cells that possess equal levels of Gag and IN. These experiments will indicate whether IN degradation is an essential feature of transposition.

C. The contribution of Tf1 RT to transposition.

To identify essential functions of RT for transposition, we have characterized a set of point mutations created either by random or site-directed mutagenesis.

1. Mutations in RNaseH that block transposition but not reverse transcription.

We have previously reported the isolation of a large set of random point mutations in Tfl that reduced transposition activity. These were screened using our homologous recombination assay to identify mutations that, never the less, produced normal levels of reverse transcript. We identified 7 Tfl mutations that exhibited these properties and as expected, most of the base changes were found to be located within IN. Surprisingly, 2 of the mutations were within the C terminus of the RNaseH domain of RT. This was an unexpected result because studies of RNaseH mutants in retroviruses have indicated that the synthesis of mature double-stranded cDNA requires RNaseH activity to release single stranded DNA intermediates from their RNA templates [Champoux, 1993 #1442; Blain, 1995 #1602]. To establish if the level of reverse transcription is effected in the two mutant strains, we extracted total DNA from each and subjected it to DNA blot analysis using conditions that allowed us to visualize restriction products from both ends of the cDNA. Consistant with the homologous recombination assays, the blot results showed that one of the strains had fully normal levels of reverse transcript while the other exhibited about 2-fold less cDNA than wild type Tfl. One possible explanation of our results is that the mutants specifically affect the ability of RNaseH to cleave the residue RNA primers off the 5' ends of the cDNA without reducing the enzyme's general ability to degrade RNA annealed to DNA. The result of this defect would be to allow RT to product completed cDNA but the presence of the RNA primers would inhibit the integration reaction. Whether or not this is the cause for the low transposition frequencies of the 2 RNaseH mutations, we appear to have identified a role of RNaseH in the integration process.

2. Site-directed mutations of Tfl RT define essential functions.

Despite the presence of highly conserved amino acid residues in the sequence of Tfl that indicate the presence of an RT, we had no direct evidence that the domain actually encodes a functional reverse transcriptase. As discussed in section A3 (*In vivo* experiments to distinguish between two types of self-priming mechanisms), we generated 8 site-directed mutations within conserved residues of Tfl RT. Within the RNaseH domain, we changed two conserved aspartic acids (443 and 498, HIV numbers) to asparagines and alanines while glu478 was changed to a glutamine. The result of each of these 5 RNaseH mutations was a Tfl element that exhibited no detectable transposition activity. In addition, the mutations resulted in undetectable levels of homologous recombination between the Tfl plasmid and cDNA, indicating the levels of reverse transcripts was low. To gauge the effect of these RNaseH mutations on the RT polymerase activity, we partially purified particles from the mutant transposons and assayed them for RT activity by adding oligo dT as primer and poly rA as template. The results indicated that D443N, D498N, D443A, and D498A all showed high levels of activity that were either equivant to wild type or within 2-fold of wild type RT. The high level of RT activity possessed by these mutant RTs indicate that the defects in transposition were not due to lower stability of RT but probably the lack of RNaseH activity. This conclusion was supported by the absence of any products of reverse transcription as measured by DNA blotting. As a result, these mutations provided the first physical data that the Tfl RT possesses RNaseH activity essential for transposition.

To address the role of the polymerase domain of RT, three additional mutations were produced in the conserved "YXDD" box that constitutes a portion of the polymerase active site in other retroelements. The mutations created were D185N, D185L, and D186N (HIV number system). The transposition frequencies of all three strains were undetectable as were their frequencies of homologous recombination. *In vitro* assays of particles produced by these mutant strains indicated the levels of RT activity was

reduced by at least 60-fold as a result of each mutation. The results from the mutations in polymerase and RNaseH strongly indicate Tfl does encode an RT that possess functions essential for transposition.

D. The characterization of host mutations that inhibit transposition.

In order to improve our understanding of the retrotransposition process and by analogy, retrovirus infection, we initiated a large scale genetic screen for *S. pombe* strains that were defective for transposition. As reported last year, cultures of *S. pombe* that contained our Tfl assay plasmid were mutagenized with EMS under conditions that caused 80% lethality on minimal medium. After these strains were colony purified, 5,000 were patched onto our transposition assay media and screened for reduced frequencies of transposition. Approximately 176 strains from two mating types reproducibly transposed at significantly lower levels than the parent strains. All these candidates were tested for the presence of single genetic lesions affecting transposition and for the magnitude of the deficiency caused by each mutation. As a result, we chose to focus our experiments on 6 mutant strains that had transposition frequencies 5 to 25-fold lower than that of wild type cells.

1. Recombination analysis of 6 mutant strains to identify the minimum number of genes that effect transposition.

Each of the original mutant strains were crossed with others from the same set and the resulting spores were tested for transposition to determine if any two strains carried mutations within the same gene. If two strains are mated that each possess mutations in the same gene, none of the resulting spores will transpose at wild type frequencies. If the mutations exist in different genes that segregate independently, one quarter of the spores will show normal transposition activity. The results indicated that each of the 6 mutant strains contained genetic lesions in different genes that we have termed *hop1* through *hop6*.

2. The evaluation of each transposition mutation for effects on specific steps in the transposition pathway.

1. *Hop1*

The strains with the *hop1-1* mutation grow at wild type rates and possess a mutation that reduced its transposition frequency by 25-fold when measured with our quantitative assay. The frequency of homologous recombination between Tfl cDNA and the Tfl plasmid was also measured quantitatively and found to be 50-fold lower in *hop1-1* cells compared to wild type levels. The low frequencies of homologous recombination suggested that the *hop1-1* mutation caused a defect early in the transposition process that resulted in lower levels of reverse transcript. Results from DNA blots of *hop1-1* strains were consistent with the recombination data in that no reverse transcripts were detected. Immunoblots of *hop1-1* strains demonstrated a dramatic lack of Gag and IN protein accumulation that very likely caused the low levels of cDNA due to a simple lack of particle formation. We have measured mRNA levels of Tfl-*neo* in *hop1-1* strains and find at most a reduction of 2-fold that may not be the main cause of the transposition defect but merely the result of transcript destabilization due to low rates of Tfl translation. We are currently pursuing the possibility that the lack of Tfl protein is due to either reduced translation of the Tfl transcript or increased degradation of the Tfl proteins. We have made lacZ fusions to the first codon of the Tfl ORF as well as to internal locations to identify which Tfl sequences are effected by the *hop1-1* defect. The isogenic *S. pombe* strains with and without the *hop1-1*

mutation are now being constructed to test the lacZ plasmids. In addition, we have fused the complete Tfl ORF to the *nmt1* promoter after the *nmt1* 5' untranslated region to determine if the *hop1-1* mutation caused a lack of Tfl protein accumulation due strictly to protein degradation and not lack of translation. Preliminary results indicate that when the Tfl proteins are expressed from the *nmt1* transcriptional leader, the *hop1-1* mutation has no effect on Gag or IN accumulation suggesting that the lack of Tfl protein accumulation caused by *hop1-1* was due to a defect in translation linked to the 5' untranslated region of the Tfl mRNA.

Because the lack of homologous recombination is so dramatic, we were able to identify plasmids from an *S. pombe* genomic library that suppressed the *hop1-1* defects. We are now characterizing these plasmid candidates to determine if they contain the *hop1* gene.

2. *hop2*

Strains with the *hop2-1* mutation have an 8-fold transposition defect and a 4-fold drop in their homologous recombination frequency. The phenotype of the *hop2* strains is similar to *hop1* cells in that the levels of Tfl protein accumulation was low although the amounts of cDNA detected by DNA blot were not greatly reduced. We are testing these strains for defects in translation and protein stability using the same lacZ fusions developed for characterizing *hop1*.

3. *hop3*

The *hop3-1* mutation caused a 10-fold drop in both the transposition and homologous recombination assay frequencies. Unlike the *hop1* and *hop2* mutations, *hop3* cells were able to produce wild type levels of mature Tfl proteins in log phase cultures. However, 1 day after *hop3-1* cells reached stationary phase, they exhibited a sharp drop in Gag levels suggesting a loss of protein stability. Although we observed the *hop3-1* cells contained normal levels of reverse transcript, the loss of Gag and possibly IN could result in lower IN activity or less efficient cDNA presentation to the nucleus. To reveal the cause of low Tfl protein accumulation in stationary phase cells we are testing these strains for defects in translation or protein stability using the same lacZ fusions developed for characterizing *hop1*. Preliminary results indicate that the Gag protein in stationary phase *hop3-1* cells is absent even when expressed from a plasmid with the *nmt1* 5' untranslated region fused to the Tfl ORF. This suggests that the lack of Gag is due to degradation, not a defect in translation.

4. *hop5*

Cells with the *hop5-1* mutation transpose with frequencies 12-fold lower than wild type strains. Results from the homologous recombination assay also indicated a 12-fold defect. As observed for *hop3-1* cells, DNA blots made from *hop5-1* strains showed normal levels of mature reverse transcript were produced. Immunoblots indicate that this mutation has no effect on the levels of Tfl protein accumulation. We speculate that the defect caused by *hop5-1* is similar to that of *hop3-1* in that a block may occur in the presentation of the cDNA:IN complex to the nucleus.

Significance to Biomedical Research and the Program of the Institute.

The medical significance of our research stems from the close relationship between retroviruses and retrotransposons. The high level of similarity in structure between these two types of retroelements

results in the large number of mechanisms that the elements share. The mechanism of IN mediated insertion for retrotransposons and retroviruses has been shown to require the same reaction intermediates [Eichinger, 1988 #419]. The early steps of reverse transcription are primed from the same position of the elements and both types of retroelements use polypurine sequences for priming plus strand reverse transcription. The structure of transposon virus-like particles is analogous to retrovirus particles in their composition of capsid and Pol proteins. The medical importance of retroviruses is dominated by the role of HIV in the AIDS epidemic. HIV is a retrovirus that appears to undergo the same types of reverse transcription, proteolysis, integration and particle formation as do retrotransposons. Because many aspects of retrovirus function, including particle assembly, reverse transcription, nuclear entry and chromosome integration, are difficult to study in hosts that possess a high level of genetic complexity, the molecular genetic analysis of retrotransposons in yeast provides a promising approach to answering many of the important questions that are at the center of the AIDS crisis.

The study of transposons in general and retrotransposons in particular is relevant to the understanding of neoplastic disease mechanisms. Transposons are endogenous mutagens of cells; chemical mutagenesis is strongly correlated with carcinogenesis; therefore studies of these endogenous biological mutagens is highly pertinent to an understanding of cancer. Moreover, the growth of a wide variety of tumors is associated with activation of endogenous retroviral expression and rearrangements of the genomic DNA in tumor cells. Many studies clearly indicate that the activation of cellular oncogenes is directly responsible for mammalian neoplasms. In many cases, such gene activation is caused by retrovirus or retrotransposon transpositional insertion into a proto-oncogene locus. For this reason, the thorough understanding to be gained from studies of retrotransposons inhabiting the genomes of genetically tractable organisms, particularly yeast, is directly relevant to the understanding of both oncogene activation mechanisms and the DNA rearrangements that occur in cancer cells.

Proposed Course

A. Self-primed reverse transcription.

1. The mechanism of mRNA cleavage.

We are pursuing several approaches to identify the chemical nature of the self-priming mechanism of Tfl reverse transcription. We have previously shown that the interaction between the PBS and the 5' end of the Tfl transcript is required for synthesis of the minus-strand strong-stop DNA. The chemical nature of the priming event is still unknown. One possibility is that a 3'OH is provided by a site-specific nicking enzyme that creates a primer by cutting off the first 11 bases of the Tfl transcript. Our data from the RT polymerase mutations has provided the first evidence for a cleavage of the Tfl transcript after the 11th base. Tfl mRNA isolated from particles with the mutant reverse transcriptases was analyzed by primer extension and found to have been cleaved between the 11th and 12th bases. We will also subject the particle RNA preparations to S1 nuclease analysis to seek independent evidence for the absence of the first 11 bases. Because the mutations in the conserved residues of RNaseH surprisingly resulted in much lower amounts of minus-strand strong-stop DNA, we will test the hypothesis that RNaseH is the source of the cleavage activity. We will compare Tfl elements with mutations in the polymerase domain to Tfl elements with mutations in both the polymerase and RNaseH domains to determine if the addition of the RNaseH mutation reduces the level of cleaved mRNA.

2. *In vitro* analysis of the self-priming mechanism.

As an additional means of studying the chemistry of priming, we will develop an *in vitro* priming assay to generate large quantities of priming intermediates. Initially, we will attempt to produce a priming reaction from purified RT and Tfl transcript. RT expression vectors will be produced and tested for use in bacteria as well as baculovirus and *Pichia pastoris*. Once RT is successfully purified from an expression system, we will combine RT, *in vitro* transcribed Tfl mRNA, and ^{32}P labeled dNTP's under a variety of conditions. These reactions will be tested for the synthesis of minus-strand strong-stop DNA. Mutations in RNaseH and/or the polymerase domain of RT will be tested in the assay to characterize the source of the cleavage activity. We will also test reactions with mRNA alone to determine if the cleavage is autocatalytic. The products of the *in vitro* reactions will also be examined for evidence of a 2'OH priming mechanism. The products of priming reactions containing radiolabeled Tfl mRNA will be digested with the nucleases P1 and T2 to determine if the bond between the RNA primer and the synthesized DNA is a 2'-5' or a 3'-5' linkage [Ruskin, 1990 #917; Ruskin, 1985 #747].

3. The function of the 39 base pair RNA structure.

In conjunction with the study of the mRNA cleavage mechanism, we will continue to explore the role of the 39 base pair RNA structure that contains the PBS. The data from our mutagenesis of this structure indicated that sequence along its entire length are critical for transposition. We now plan to evaluate if these mutations can be rescued by the addition of compensatory mutations, suggesting that the RNA structure is important and not the individual sequences. Mutations from each of the 4 sections of the RNA structure will be further studied to identify their individual contributions to the transposition process. These mutations will be tested for their effects on translation, mRNA packaging into particles, self-priming, reverse transcription, and integration.

B. The mechanism of Tfl VLP assembly.

1. Identification of growth conditions that induce IN degradation

We have shown that cells starved for nitrogen or sulfate undergo IN degradation while cultures starved for glucose maintain equal molar amounts of Gag and IN. This observation will allow us to test the transposition efficiency of Tfl in cells that starve for glucose to evaluate the importance of IN degradation to transposition. If functional particle development requires IN degradation then the cells starved for glucose will be unable to transpose even several days after stationary phase is reached. If however, the degradation of IN is not required for transposition, the glucose starved cells should undergo normal levels of transposition. This question of IN degradation will be addressed below in separate experiments.

2. The role of IN degradation in transposition.

Although we propose that the reduction in mature IN protein observed in stationary phase cells is a regulated step of Tfl particle assembly, the possibility exists that assembly occurs before IN degradation and excess IN not incorporated into the VLP's is degraded simply due to greater exposure to cellular proteases. Our observation that much of Tfl reverse transcription occurred after IN degradation indicated that the particles with excess Gag are functional and represent intermediates in the transposition process. The idea that IN degradation is a programmed aspect of particle development is consistent with our observation that the majority of the RT is also degraded in stationary phase cells

(J. Lin, A. Atwood, and H. Levin, unpublished). A resolution to the question of how important is the reduction of Pol protein to the assembly process requires that we develop direct genetic means of manipulating the Gag to Pol ratios in stationary phase cells.

We plan to test the importance of high Gag to IN ratios in transposition by overexpressing IN from a separate plasmid with a promoter that is active in stationary phase cells. If particle development requires IN levels be reduced much below the amount of Gag protein, the increase in IN provided by the overexpression plasmid should cause a defect in transposition. The *fbp1* promoter is active in stationary phase and will be tested for use in this experiment [Hoffman, 1989 #858].

Another approach we will take to define the importance of Gag to IN ratios is to reduce Pol protein levels throughout particle formation by inserting "leaky" frameshift mutations between Tfl Gag and PR, PR and RT, and between RT and IN. The frameshift sequences from Ty3, Ty1 and M1 have been characterized in *S. cerevisiae* and will be tested in *S. pombe* for the purposes of this experiment [Dinman, 1992 #1361; Farabaugh, 1993 #1389; Belcourt, 1990 #824]. If the excess Pol proteins are simply an impediment to assembly, the reduction in expression caused by the frameshift mutations will have no detrimental effect on transposition and may actually cause an improvement in activity.

3. Characterization of the activity that degrades IN.

An important aspect of Tfl particle assembly that we will address here is the source of the IN degradation activity. The protease or proteases that degrade the IN and RT in stationary phase cells could either be a transposon encoded protein such as PR, or it could be a host protein. We will study the stability of mature-sized IN expressed on a plasmid in the absence of all the other Tfl proteins to determine if host factors alone are sufficient to regulate the IN degradation in stationary phase cells without the association of other Tfl proteins. If no IN degradation is observed in these stationary phase cells, or if IN is unstable even in log-phase cells, the possibility will be tested that regulated degradation of IN requires the presence of Gag, RT, and IN assembled into particles. We plan to construct Gag, PR, and RT expression plasmids so that mature-sized Gag, PR, RT, and IN can be independently expressed within the same cells to evaluate the requirements for IN degradation. IN and PR will be coexpressed in one experiment as a direct test of PR catalyzed degradation of IN. In additional experiments, Gag, PR, RT and IN will be coexpressed in several combinations to establish the minimal requirements to reproduce the regulated degradation of IN. If IN regulation can be observed in one of these combinations, the requirement for PR will be evaluated.

The independent expression of all of the mature Tfl proteins will present an opportunity to determine if particle assembly requires intermediates composed of primary translation products. One model for particle assembly predicts that the attachment of Gag to the Pol proteins in the primary translation product may be required to localize RT and IN to the sites of particle formation. Alternatively, the result that Tfl proteins are completely processed into mature products that appear to form particles before RT and IN are degraded suggests that the independent expression of Gag, RT and IN may result in functional particles. The ability of independently expressed mature-sized proteins to form functional particles will be tested by expressing Gag from a Tfl-*neo* transcript that has a frame shift placed before PR.

4. The study of particle stability to measure the structural integrity of log phase particles compared to stationary phase particles.

The Tfl protein isolated from stationary phase cells was found to sediment in sucrose gradients as particles that contained 50-fold more Gag than IN [Atwood, 1995 #1576]. Particles isolated from log phase cells have very different properties including the association of a much greater amount of IN and RT [Atwood, 1995 #1576]. We wish to probe these two forms of Tfl particles for evidence of a maturation process that could result in greater structural integrity. The stability of the log and stationary phase particles will be evaluated by their resistance to perturbations such as treatment with denaturing agents and proteases. Particles from log and stationary phase cells will be isolated in sucrose gradients. The strength of the Gag and IN association in these particles will be tested in various concentrations of denaturing agents (e.g. tween, urea and guanidinium hydrochloride) to evaluate the stability of log-phase particles compared to stationary phase particles. After the denaturing treatments, the particles will be refractionated on sucrose gradients and the fractions analyzed on immunoblots to determine the effect of the treatments on particle integrity. The structure of log phase and stationary phase particles will also be probed by partial protease treatments. The sensitivity of Gag and IN to proteases will also be monitored by immunoblot analysis.

C. The primer cleavage activity of RNaseH.

Two different mutations in the RNaseH domain of RT were identified that caused a sharp drop in transposition activity despite the results of DNA blot analysis that showed these strains produced normal levels of full-length reverse transcript. This was an unexpected result because previous studies of RNaseH mutants have indicated that the synthesis of mature double-stranded cDNA requires RNaseH activity to release single stranded DNA intermediates from their RNA templates [Champoux, 1993 #1442; Blain, 1995 #1602]. One likely explanation of these results is that the mutants specifically affect the ability of RNaseH to cleave the residue RNA primers off the 5' ends of the cDNA without reducing the enzyme's general ability to degrade RNA annealed to DNA. The result of this defect would be to allow RT to product completed cDNA but the presence of the RNA primers would inhibit the integration reaction. Although this primer cleavage activity has been observed *in vitro* [Champoux, 1984 #1603; Rattray, 1987 #1606; Furfine, 1991 #1604; Pullen, 1992 #1605], no evidence has yet been reported indicating the two RNaseH activities have different active site requirements. We will test this hypothesis using a primer extension method and a DNA blot technique to detect the cDNA primers. If the mutants exhibit normal primer removal, we will consider alternative causes of the reduced transposition such as participation of RNaseH in the integration reaction. This issue could be addressed using the type of *in vitro* integration assays developed for other retrotransposons [Kirchner, 1995 #1474; Braiterman, 1994 #1492; Braiterman, 1994 #1493].

D. The characterization of host gene function in Tfl retrotransposition.

1. *hop1*.

The observation that *hop1-1* mutant cells do not accumulate Tfl protein could be due to a defect in translation or an increase in degradation. We have now assembled a set of Tfl expression plasmids that contain lacZ fused to several different regions of Tfl. These plasmids will be expressed in *S. pombe* to determine whether the Tfl 5' untranslated region or the Tfl coding sequence must be fused to lacZ for beta-galactosidase activity to be reduced in *hop1-1* cells. We have also constructed a plasmid containing the Tfl ORF fused to the *nmt1* promoter with the *nmt1* 5' untranslated region. This plasmid will be transformed into wild type and *hop1-1* cells to test the stability of normal Tfl proteins expressed from this heterologous transcript. Any effect of the *hop1-1* mutant on Gag levels expressed from this

plasmid would suggest that Tfl protein levels drop due to increased degradation. The results produced using all the fusion plasmids will allow additional plasmids to be produced that will help to further define which sequences of Tfl are most effected by *hop1-1*.

We will seek additional information about the function of *hop1* by continuing our effort to clone the gene. We now have one 6 kb genomic clone that suppresses the *hop1-1* phenotype and is genetically linked to the site of the original *hop1* mutation. Should this 6 kb insert contain the *hop1* gene, we will analyze its sequence for information related to its function.

2. *hop2*.

Strains with the *hop2-1* mutation have an 8-fold transposition defect and a 4-fold drop in their homologous recombination frequency. Because the phenotype of the *hop2* strains is similar to *hop1* cells in that the levels of Tfl protein accumulation was low, we are testing these strains for defects in translation or protein stability using the same *lacZ* fusions developed for characterizing *hop1*. We will also attempt to isolate the *hop2* gene using an *S. pombe* genomic library to identify complementing clones.

3. *hop3*.

The *hop3-1* mutation caused a 10-fold drop in both the transposition and homologous recombination assay frequencies. Unlike the *hop1* and *hop2* mutations, *hop3* cells were able to produce wild type levels of mature Tfl proteins in log phase cultures. However, 1 day after *hop3-1* cells reached stationary phase, they exhibited a sharp drop in Gag levels suggesting a loss of protein stability. Although we observed the *hop3-1* cells contained normal levels of reverse transcript, the loss of Gag and possibly IN could result in lower IN activity or less efficient cDNA presentation to the nucleus. To reveal the cause of low Tfl protein accumulation in stationary phase cells we are testing these strains for defects in translation or protein stability using the same *lacZ* fusions developed for characterizing *hop1*. Preliminary results indicate that the Gag protein in stationary phase *hop3-1* cells is absent even when expressed from a plasmid with the *nmt1* 5' untranslated region fused to the Tfl ORF. This suggests that the lack of Gag is due to degradation, not a defect in translation. We will therefore attempt to identify which Tfl proteins are specifically destabilized in *hop3-1* cells. We also plan to isolate the *hop3* gene using the genomic library to complement the lack of transposition exhibited by *hop3-1* cells.

4. *hop5*

Cells with the *hop5-1* mutation transpose with frequencies 12-fold lower than wild type strains. Results from the homologous recombination assay also indicated a 12-fold defect. As observed for *hop3-1* cells, DNA blots made from *hop5-1* strains showed normal levels of mature reverse transcript were produced. Immunoblots indicated that this mutation had no effect on the levels of Tfl protein accumulation. We speculate that the defect caused by *hop5-1* may occur in the presentation of the cDNA:IN complex to the nucleus. Alternatively, the transposition and recombination defect could be caused by a mutation in a general recombination factor that contributes to both processes. To test this possibility, we will measure the efficiency of homologous recombination between plasmid and genomic copies of *ade6* in both wild type and *hop5-1* cells. We will also measure the frequency of switching at the mating type cassette as an independent measure of a homologous recombination process.

Additional information about the function of *hop5* will be sought by the isolation and characterization of its sequence.

Publications

Atwood A, Lin J, Levin H. The retrotransposon Tfl assembles virus-like particles with excess Gag due to specific degradation of Pol protein, *Mol Cell Biol* 1995; in press.

Levin HL. A novel mechanism of self-primed reverse transcription defines a new family of retroelements, *Mol Cell Biol* 1995;15:3310-17.

References

- Atwood, A., J. Lin, et al. (1995). "The retrotransposon Tfl assembles virus-like particles with excess Gag due to specific degradation of Pol protein." *Mol. Cell. Biol.*: in press.
- Belcourt, M. F. and P. J. Farabaugh (1990). "Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site." *Cell* **62**: 339-352.
- Blain, S. and S. Goff (1995). "Effects on DNA synthesis and translocation caused by mutations in the RNase H domain of Moloney Murine Leukemia virus reverse transcriptase." *J. Virol* **69**: 4440-4452.
- Braiterman, L. T. and J. D. Boeke (1994). "In vitro integration of retrotransposon Ty1: a direct physical assay." *Mol Cell Biol* **14**(9): 5719-30.
- Braiterman, L. T. and J. D. Boeke (1994). "Ty1 in vitro integration: effects of mutations in cis and in trans." *Mol Cell Biol* **14**(9): 5731-40.
- Champoux, J. (1993). Roles of ribonuclease H in reverse transcription. *Reverse Transcriptase*. A. Skulka and S. Goff. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press: 103-117.
- Champoux, J., E. Gilboa, et al. (1984). "Mechanism of RNA primer removal by the RNase H activity of avian myeloblastosis virus reverse transcriptase." *J. Virol* **49**: 686-691.
- Davies, J., Z. Hostomska, et al. (1991). "Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase." *Science* **252**: 88-95.
- Dinman, J. D. and R. B. Wickner (1992). "Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation." *Journal of Virology* **66**: 3669-3676.
- Eichinger, D. J. and J. D. Boeke (1988). "The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition." *Cell* **54**(955): 955-966.
- Farabaugh, P. J., H. Zhao, et al. (1993). "A novel programmed frameshift expresses the POL3 gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage [published erratum appears in *Cell* 1993 Nov 19;75(4):826]." *Cell* **74**(1): 93-103.
- Furfine, E. and J. Reardon (1991). "Reverse transcriptase-RNase H from the human immunodeficiency virus. Relationship of the DNA polymerase and RNA hydrolysis activities." *J Biol Chem* **266**: 406-412.
- Hoffman, C. S. and F. Winston (1989). "A transcriptionally regulated expression vector for the fission yeast *Schizosaccharomyces pombe*." *Gene* **84**: 473-479.
- Hostomsky, Z., S. H. Hughes, et al. (1994). "Redesignation of the RNase D activity associated with retroviral reverse transcriptase as RNase H." *J Virol* **68**(3): 1970-1.
- Hsu, M., S. Eagle, et al. (1992). "Cell-free synthesis of the branched RNA-linked msDNA from retron-Ec67 of *Escherichia coli*." *J. Biol. Chem.* **267**: 13823-13829.

- Hsu, M. Y., M. Inouye, et al. (1990). "The msDNA-Ec67 retron: a potential *E. coli* transposable element encoding both reverse transcriptase and Dam methylase proteins." Proc. Natl. Acad. Sci. USA **87**: 9454-9458.
- Katayanagi, K., M. Miyagawa, et al. (1990). "Three dimensional structure of ribonuclease H from *E. coli*." Nature **347**: 306-309.
- Kirchner, J., C. Connolly, et al. (1995). "Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element." Science **267**: 1488-1491.
- Kohlstaedt, L., J. Wang, et al. (1992). "Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor." Science **256**: 1783-1790.
- Lampson, B. C., M. Inouye, et al. (1989). "Reverse transcriptase with concomitant ribonuclease H activity in the cell-free synthesis of branched RNA-linked msDNA of *Myxococcus xanthus*." Cell **56**: 701-707.
- Levin, H. L. (1995). "A Novel mechanism of self-primed reverse transcription defines a new family of retroelements." Mol Cell Biol **15**: 3310-3317.
- Levin, H. L. and J. D. Boeke (1992). "Demonstration of retrotransposition of the Tfl element in fission yeast." EMBO J. **11**: 1145-1153.
- Mizrahi, V., R. Brooksbank, et al. (1994). "Mutagenesis of the conserved aspartic acid 443, glutamic acid 478, and aspartic acid 498 residues in the ribonuclease H domain of p66/p51 human immunodeficiency virus type I reverse transcriptase." Journal of Biological Chemistry **269**: 19245-19249.
- Moreno, S., A. Klar, et al. (1991). "Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*." Methods in Enzymology **194**: 795-823.
- Panet, A., D. Baltimore, et al. (1975). "Quantitation of avian RNA tumor virus reverse transcriptase by radioimmunoassay." Journal of Virology **16**: 146-152.
- Pullen, K., L. Ishimoto, et al. (1992). "Incomplete removal of the RNA primer for minus-strand DNA synthesis by human immunodeficiency virus type I reverse transcriptase." J Virol **66**: 367-373.
- Ratray, A. and J. Champoux (1987). "The role of Moloney murine leukemia virus RNase H activity in the formation of plus-strand primers." J Virol **61**: 2843-2851.
- Repaske, R., J. Hartley, et al. (1989). "Inhibition of RNase H activity and viral replication by single mutations in the 3' region of moloney murine leukemia virus reverse transcriptase." Journal of Virology: 1460-1464.
- Ruskin, B. and M. R. Green (1985). "An RNA processing activity that debranches RNA lariats." Science **229**: 135-140.
- Ruskin, B. and M. R. Green (1990). "RNA lariat debranching enzyme as tool for analyzing RNA structure." Meth. Enzymol. **181**: 180-188.
- Stromberg, K., N. Hurley, et al. (1974). "Structural studies of avian myeloblastosis virus: Comparison of polypeptides in virion and core components by dodecyl sulfate-polyacrylamide gel electrophoresis." Journal of Virology **13**: 513-528.
- Yang, W., W. Hendrickson, et al. (1990). "Structure of ribonuclease H phased at 2 Å resolution by MAD analysis of the selenomethionine protein." Science **249**: 1398-1405.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01-HD01010-01

PERIOD COVERED

October 1, 1994 through September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Regulation of Eukaryotic Protein Synthesis

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	Thomas E. Dever	Unit Head	LMG:NICHD
Others:	Kobayashi, Makiko	Visiting Fellow	LMG:NICHD
	Locke, Emily	Biologist	LMG:NICHD
	Ung, Tekly	Volunteer	LMG:NICHD

COOPERATING UNITS (if any)

LAB/BRANCH

Laboratory of Molecular Genetics

SECTION

Unit on Translational Regulation

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

3.4

PROFESSIONAL:

1.9

OTHER:

1.5

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unrounded type. Do not exceed the space provided.)

The regulation of protein synthesis in mammalian cells under stress conditions as well as the regulation of GCN4-specific translation in the yeast Saccharomyces cerevisiae under amino acid starvation conditions is mediated by phosphorylation of the eukaryotic translation initiation factor (eIF)-2. Phosphorylation of serine-51 on the α subunit of eIF-2 by the mammalian PKR kinase inhibits general translation, while phosphorylation of serine-51 by the yeast GCN2 kinase is required to increase expression of GCN4. We have been studying how these protein kinases specifically recognize and phosphorylate eIF-2 α on serine-51. Over 100 mutations throughout the eIF-2 α protein have been identified that prevent induction of GCN4 expression when GCN2 is activated. Among the residues immediately flanking serine-51, the positions -1 and -2 appear most important for regulation. Examination of eIF-2 α phosphorylation in vivo in strains carrying various mutant alleles reveals that some of the substitutions markedly affect the ability to phosphorylate serine-51. We have also been studying the vaccinia virus K3L protein, a pseudosubstrate inhibitor of PKR. We have established a system to suppress the toxicity of PKR expression in yeast by co-expressing K3L. Using this system both loss-of-function and increased activity K3L mutants have been isolated. Residues near the carboxyl-terminus of K3L, that are conserved in eIF-2 α (residues 73-83), are critical for K3L activity. In addition, the corresponding residues in eIF-2 α are required for proper regulation of GCN4 expression. Finally, we have begun characterization of a novel eIF-2 α kinase inhibitor from the baculovirus Autographa californica.

Project Description:

Objective: To understand the mechanism and regulation of protein biosynthesis in eukaryotic organisms, especially focusing on the phosphorylation of eIF-2 α by the yeast *Saccharomyces cerevisiae* GCN2 kinase and by the mammalian PKR kinase, and to understand how viruses circumvent this regulatory network.

Major Findings:1. Identification of eIF-2 α mutants defective for GCN4 translational regulation.

The yeast GCN2 kinase phosphorylates eIF-2 α on serine-51 in response to an amino acid starvation. Phosphorylated eIF-2 acts as a competitive inhibitor of eIF-2B, the guanine nucleotide exchange factor for eIF-2. Thus, phosphorylation of eIF-2 results in an inhibition of general translation. In yeast, this phosphorylation of eIF-2 and subsequent inhibition of eIF-2B activity not only limits general translation, but also specifically stimulates *GCN4* expression. Increased expression of GCN4 is required for yeast strains to grow under amino acid starvation conditions. Therefore, yeast mutants that are unable to phosphorylate eIF-2 α or which fail to inhibit eIF-2B cannot induce *GCN4* expression and are thus sensitive to amino acid starvation conditions. To identify how the eIF-2 α kinases, such as GCN2, specifically recognize serine-51 on the α subunit of eIF-2 we have begun a mutagenic analysis of eIF-2 α . The residues immediately flanking serine-51 were randomly mutated and the mutant alleles were introduced into yeast in place of the wild-type eIF-2 α . Pools of mutants were screened to identify eIF-2 α mutations that block the growth of the yeast cells under amino acid starvation conditions. Following the identification of the mutants sensitive to starvation conditions, the plasmids carrying the mutant eIF-2 α alleles were isolated and sequenced to identify the amino acid substitution in eIF-2 α that blocks the translational regulation of *GCN4* expression. The sequence flanking serine-51 from residue 48 (-3) to residue 54 (+3) is: SELS₅₁RRR. No loss-of-regulation mutations were identified at residues S48, R52 or R53. Only a single mutation was identified at the +3 position: R54G. However, 10 mutations at position -1 (L50) and all 19 possible substitutions at position -2 (E49) blocked the regulation. There are at least two possible mechanisms by which a mutation in eIF-2 α could block regulation of *GCN4* expression: (1) the mutation could block the ability of GCN2 to phosphorylate eIF-2 α or (2) the mutation could prevent phosphorylated eIF-2 from regulating eIF-2B. One-dimensional polyacrylamide isoelectric focusing slab gels can be used to examine eIF-2 α phosphorylation in yeast cells. To date, only one of the loss-of-regulation mutations, a leucine to proline change at position 50 (L50P), has been found to severely inhibit phosphorylation of eIF-2 α by GCN2.

In addition to examining the residues immediately flanking serine-51, we have also mutated other residues based on their conservation in the vaccinia virus K3L protein. The vaccinia virus K3L protein is an 88 amino acid pseudosubstrate inhibitor of the mammalian anti-viral PKR kinase. K3L shares 28% amino acid sequence identity with the amino-terminal 90 residues of eIF-2 α . This homology is most striking between residues 72 and 83 of eIF-2 α where 10 of 12 residues are conserved between K3L and human eIF-2 α . Based on the strong sequence conservation and the fact that both K3L and eIF-2 α bind to PKR we hypothesized that the conserved residues may be important for the kinases to recognize and phosphorylate eIF-2 α . To test this prediction the residues in yeast eIF-2 α that are conserved in K3L were randomly mutated and the mutant pools screened to identify eIF-2 α alleles that block regulation. To date, 14 residues have been examined and 72 loss-of-regulation mutations have been identified. It appears that for several of the residues in the conserved block between positions 72 and 83 any mutation will obstruct regulation. Preliminary analysis of several of these mutations using the isoelectric focusing gels suggests that they do not impede the ability of GCN2 to phosphorylate eIF-2 α . Therefore, these mutations are probably affecting the inhibition of eIF-2B by phosphorylated eIF-2. Since both K3L and eIF-2 α bind to the eIF-2 α kinases

it seems reasonable to suggest that the residues conserved between K3L and eIF-2 α are probably important for this interaction. However, the mutational and isoelectric focusing analysis of some of these conserved residues suggests that they are affecting the interaction between eIF-2 and eIF-2B. Taken together these results suggest a model in which the kinase recognition and eIF-2B recognition surfaces on eIF-2 α overlap.

2. Mutational analysis of the vaccinia virus K3L protein.

As introduced above, the vaccinia virus K3L protein is a pseudosubstrate inhibitor of the eIF-2 α kinase PKR, a component in the mammalian antiviral defense mechanism. To increase our understanding of how the eIF-2 α kinases recognize their substrate we have carried out an analysis of the K3L protein. First, we obtained from Dr. Patrick Romano a yeast strain in which the PKR kinase, expressed under the control of a GAL promoter, was integrated into the yeast genome. This strain grows well on glucose medium, however the strain is dead on galactose medium due to severe inhibition of translation caused by phosphorylation of eIF-2 α . The vaccinia virus K3L protein gene was cloned into the vector pEMBLyex4 such that K3L expression was also under GAL control. When the K3L expression plasmid was introduced into the yeast strain expressing PKR we found that K3L could suppress the lethality caused by high level expression of PKR in yeast. These results established the yeast system as a tool to study K3L.

Two types of mutational analyses have been conducted on K3L: a random mutational analysis to identify superactive K3L alleles and a site-directed mutational analysis to assess the importance of residues conserved between K3L and eIF-2 α . For the random mutational analysis the K3L gene was amplified under low-fidelity conditions using the PCR. The mutant pool of K3L genes was subcloned into pEMBLyex4 and the plasmids were introduced into the yeast strain containing the GAL-PKR construct. The colonies were screened to identify K3L mutants that were better suppressors of PKR. Five such mutants have been identified and one, which is significantly better than the others, has been more extensively characterized. This latter K3L upmutant gene was sequenced and found to contain two nucleotide substitutions. One was a silent mutation that did not alter the amino acid sequence, and the second caused a histidine to arginine change at residue 47 (K3L-H47R). While wild-type K3L yielded a partial suppression of the slow-growth phenotype due to high level expression of PKR in yeast, the K3L-H47R mutant afforded almost complete suppression of the toxic effects associated with PKR. Immunoblot analysis reveals that the K3L-H47R protein is not expressed to higher levels than the wild-type protein nor is PKR expression lower in the K3L-H47R strain than in the wild-type K3L strain. Thus the increased suppression is not due to trivial reasons like increased K3L expression or decreased PKR levels. One hypothesis we are currently pursuing is that K3L-H47R will bind tighter to PKR than wild-type K3L. We have been able to co-immunoprecipitate K3L and K3L-H47R with PKR. Currently, we are altering the wash conditions to see if we can detect a difference in the strength of the association between PKR and K3L-H47R versus wild-type K3L. A point of interest relating back to the first project is that the K3L-H47R mutation makes K3L more similar to eIF-2 α . In fact, the H47 residue in K3L aligns with R52 in eIF-2 α immediately adjacent to the phosphorylation site at serine-51. A possible interpretation of these results is that by increasing the similarity between K3L and eIF-2 α around the serine-51 region, K3L will interact more avidly with PKR.

The site-directed mutational analysis of K3L has focused on the carboxyl-terminal region. As stated above, this region of K3L contains a stretch where 10 out of 12 residues are perfectly identical to mammalian eIF-2 α . The sequence KGYID found between residues 74 and 78 in K3L is perfectly conserved in all eIF-2 α and K3L proteins that have been identified. As mentioned previously mutations in this region of yeast eIF-2 α block the regulation of GCN4 expression. Mutation of K74, Y76 or D78 of K3L to alanine blocks the ability of either wild-type K3L or the upmutant K3L-H47R to suppress PKR. In addition, immunoblot analysis reveals that the Y76A mutation does not alter K3L protein levels ruling out the trivial possibility

that the loss of suppression is due to lower K3L protein levels. A second set of site-directed mutations were introduced to determine the carboxyl-terminal boundary of K3L required for its PKR inhibitory activity. Truncation of the 88-residue K3L protein at residue 73, eliminated the KGYID sequence and the ability to suppress PKR. Truncation of only the last five residues had no effect on K3L activity. While after truncation of three additional residues (8 total), the ability of K3L or K3L-H47R to inhibit PKR was lost. Recent results reveal that after removal of the carboxyl-terminal 6 amino acids K3L-H47R retains its suppressing activity suggesting that the carboxyl-terminal boundary for K3L activity is either 6 or 7 residues from the end of the protein.

3. Identification of a baculovirus inhibitor of eIF-2 α kinases.

Many viruses have evolved ways to counteract the antiviral defense mechanisms of eukaryotic cells. One of the defense mechanisms cells employ is phosphorylation of eIF-2 α by the PKR kinase. As described above, vaccinia virus expresses the K3L protein, a pseudosubstrate inhibitor of PKR. Other viruses express different protein or RNA inhibitors of PKR and still others activate latent cellular inhibitors of PKR or even degrade PKR. Since our studies of the K3L protein have provided insights into how PKR recognizes eIF-2 α , we have examined the literature closely to see if we could identify any new viral inhibitors of PKR that might offer new insights into this mechanism of translational regulation. When the sequence of the baculovirus *Autographa californica* was reported late in 1994 an open reading frame designated ORF123 encoding a truncated protein kinase termed pk2 was identified. The amino acid sequence of pk2, though truncated, was most similar to the eIF-2 α kinase family. Since many viruses encode inhibitors of the eIF-2 α kinases, we reasoned that pk2 might be an inhibitor of PKR. To test this hypothesis the pk2 gene was isolated from baculovirus using the PCR and the gene was inserted into the vector pEMBLyex4. Introduction of the pk2 expression plasmid into the yeast strain that expresses PKR under the control of the GAL promoter revealed that pk2 could suppress the toxicity of PKR in yeast. Not only could pk2 suppress PKR toxicity, but pk2 was also able to suppress both wild type and hyperactive alleles of the yeast eIF-2 α kinase GCN2. These results demonstrate that pk2 is an eIF-2 α kinase inhibitor, and since baculovirus is an insect virus this is the first indication that insects may also use phosphorylation of eIF-2 α as a means to block viral infection.

The pk2 protein is similar to the carboxyl-terminal 1/2 of an eIF-2 α kinase domain. This raised the possibility that the corresponding fragment from an authentic eIF-2 α kinase domain may also act in a dominant negative manner. The amino acid sequences from GCN2 and pk2 were aligned and oligonucleotides were identified to amplify the corresponding region of the GCN2 kinase using the PCR. Following cloning into pEMBLyex4, the truncated gcn2 kinase domain was introduced into various strains. High level expression of the truncated gcn2 kinase domain was found to inhibit the toxicity associated with the hyperactive GCN2^c kinases and thus reveal that this region of an eIF-2 α kinase can act in a dominant negative manner.

4. Examination of substrate specificity of the eIF-2 α kinases.

Most protein kinases can be classified into one of two classes either serine/threonine specificity or tyrosine specificity. However, a few kinases have been identified with dual specificity that will phosphorylate both tyrosine and serine/threonine residues. To date the only known physiologically relevant substrate of the eIF-2 α kinases is eIF-2 α , however recent reports suggest that PKR may have alternative substrates. While the eIF-2 α kinases have only been reported to phosphorylate serine residues the alternative

phospho-acceptors threonine and tyrosine have not been examined. Intriguingly, the mouse PKR kinase was originally identified as TIK in a screen for tyrosine kinases. However, in this study there was no evidence that PKR was phosphorylated on or that PKR could phosphorylate other proteins on tyrosine residues. To further investigate the substrate specificities of the eIF-2 α kinases and as a possible means to examine the differences between the serine/threonine and tyrosine kinases we mutated the serine-51 residue in eIF-2 α to threonine and tyrosine. Yeast cells expressing the eIF-2 α -S51T allele were practically indistinguishable from wild-type cells with a good growth rate and the ability to grow under amino acid starvation conditions, thus indicating that GCN2 can phosphorylate the threonine at residue-51 on the eIF-2 α -S51T protein. In addition, high level expression of PKR in the eIF-2 α -S51T strain was lethal again demonstrating that the eIF-2 α kinases can phosphorylate a threonine residue, and that the phosphothreonine residue maintains the normal regulation of eIF-2B.

When serine-51 was replaced with tyrosine initial results suggested that tyrosine was not a substrate. Yeast strains expressing GCN2 and carrying the eIF-2 α -S51Y allele are sensitive to amino acid starvation conditions suggesting that GCN2 cannot phosphorylate eIF-2 α -S51Y. In addition high level expression of PKR in the eIF-2 α -S51Y strain is not lethal. Both of these results are consistent with the eIF-2 α kinases being unable to phosphorylate tyrosine residues. However, upon closer inspection a very modest slow-growth phenotype was observed in eIF-2 α -S51Y strains expressing high levels of PKR. In addition, these strains expressing PKR to high levels could grow under amino acid starvation conditions suggesting that PKR may be phosphorylating the tyrosine residue and inducing GCN4 expression. To provide convincing evidence that PKR can in fact phosphorylate eIF-2 α -S51Y, isoelectric focusing analysis was performed. A signal consistent with phosphorylation of eIF-2 α was observed in strains expressing PKR and either wild-type eIF-2 α or eIF-2 α -S51Y, but not eIF-2 α -S51A. In addition, the phosphorylation of eIF-2 α -S51Y was dependent on wild-type PKR and was not observed with an inactive mutant of PKR. To confirm that the eIF-2 α -S51Y protein was indeed phosphorylated on tyrosine the immunoblot from the isoelectric focusing gel was probed with anti-phosphotyrosine antibodies. A signal was specifically detected in the strains expressing both PKR and eIF-2 α -S51Y, but not wild-type eIF-2 α . Thus PKR can phosphorylate a tyrosine residue in vivo. Surprisingly, GCN2^c kinases were also found to phosphorylate eIF-2 α -S51Y when examined on the isoelectric focusing gels despite the lack of any noticeable phenotypes. The finding that PKR can phosphorylate a tyrosine residue in vivo raises the possibility that alternate substrates for PKR or any of the eIF-2 α kinases may be phosphorylated on tyrosine residues rather than exclusively on serine residues.

Proposed Course of Project:

1. eIF-2 α mutants defective for translational regulation. We plan to continue our analysis of eIF-2 α mutants identifying alleles that fail to properly induce GCN4 expression under amino acid starvation conditions. More importantly, we will examine the phosphorylation of serine-51 in the mutant proteins using the isoelectric focusing gels. While only a single mutation has been identified that severely impairs phosphorylation, a few mutations have been identified that appear to reduce phosphorylation. By constructing chimeric eIF-2 α proteins containing combinations of mutations that alone modestly impair phosphorylation, we hope to obtain additional alleles that are severely impaired for phosphorylation by GCN2 in vivo. The eIF-2 α alleles that are defective for phosphorylation by GCN2 in vivo will also be examined both in vivo and in vitro with PKR. The ability of these alleles to suppress the toxicity of high level PKR expression in yeast will be monitored, and the ability of PKR to phosphorylate these proteins will be examined both in vivo (using the isoelectric focusing gels) and in vitro. The in vitro kinase assays will employ either recombinant PKR or PKR immunoprecipitated from yeast and the substrate will be

recombinant eIF-2 α alleles expressed using bacterial expression systems. Alleles of eIF-2 α that are defective for phosphorylation by PKR will be used in a reversion analysis to identify PKR mutants with altered substrate specificity. Pools of mutant PKR genes will be introduced into yeast cells expressing the mutant eIF-2 α alleles, and PKR mutants that restore phosphorylation will be identified as causing a slow-growth phenotype or increased resistance to amino acid starvation conditions.

2. Analysis of the vaccinia virus K3L protein. The first priority is to determine if the superactive K3L mutants bind to PKR with greater avidity. As discussed above, K3L can be co-immunoprecipitated with PKR and we are now varying conditions to see if K3L-H47R binds tighter than wild-type K3L. Similarly, the loss-of-function mutants in K3L will also be examined using the co-immunoprecipitation assay to determine if these mutations alter the binding affinity of K3L for PKR. A new screen for additional K3L upmutants will be performed. Through the identification of key residues in the K3L protein we hope to uncover the recognition elements the kinases use when modifying eIF-2 α . In a second screen we will identify PKR mutants that are resistant to K3L. The K3L upmutant K3L-H47R will be introduced into a yeast strain under the control of a GAL promoter. Wild-type PKR will be non-toxic in this strain due to inhibition by K3L. The PKR gene also under GAL control will be randomly mutated and a pool of mutant PKR genes will be introduced into the yeast strain expressing K3L-H47R. PKR mutants resistant to K3L-H47R will be identified as causing a slow-growth or no-growth phenotype on galactose medium. These PKR mutants will also be examined to determine their resistance to the various K3L upmutants uncovered in the other screens. The identification of PKR mutants resistant to K3L will help identify the substrate binding region in PKR; and through the examination of the various K3L and PKR mutants we may be able to map the interacting surfaces in PKR and K3L. These results should help identify how PKR and the other eIF-2 α kinases interact with the authentic substrate eIF-2 α .

3. Analysis of the baculovirus pk2 protein. A primary goal for this study is to determine how pk2 inhibits the eIF-2 α kinases. First, we will examine eIF-2 α phosphorylation in yeast cells expressing pk2 and various eIF-2 α kinases. Based on preliminary results we expect pk2 to lower eIF-2 α phosphorylation levels. Two models can be proposed for the suppressing activity of pk2: (1) pk2 binds to eIF-2 and prevents the eIF-2 α kinases from gaining access to their substrate, or (2) pk2 heterodimerizes with the eIF-2 α kinases and the heterodimers are less active. According to the first model it would be expected that overexpressing eIF-2 or eIF-2 α alone would reverse the effects of pk2. While overexpression of GCN2 or another eIF-2 α kinase would be expected to counteract pk2 according to the second model. High level expression of pk2 causes a slow-growth phenotype in yeast cells independent of any eIF-2 α kinases. The cause of this slow-growth phenotype is unknown, however suppressors of the slow-growth may provide insights into proteins which interact with pk2. To identify dosage-dependent suppressors the pk2 gene under the control of the GAL promoter will be introduced into a yeast cell. This strain will subsequently be transformed with a high copy-number library of yeast genes and fast growing transformants will be identified. Prior to this random screen for dosage-dependent suppressors, various candidate genes (for example, eIF-2 α and GCN2) on high copy number vectors will be tested. While it is not known if the slow-growth phenotype caused by high-level expression of pk2 is related to the ability of pk2 to suppress the eIF-2 α kinases, the identification of proteins that interact with pk2 is of interest since these proteins may also interact with the eIF-2 α kinases. Another means to identify proteins that interact with pk2 is to perform co-immunoprecipitation reactions and examine if candidate proteins such as eIF-2 can be co-precipitated with pk2. To facilitate this analysis we have engineered a c-myc epitope tag on the carboxyl-terminus of pk2.

In addition to examining authentic pk2, we will continue our examination of the related domain from the eIF-2 α kinases. As reported above, the corresponding region of the GCN2 kinase will act to suppress at least two different GCN2^c kinases. Similar regions of PKR and HRI have also been examined, however

these proteins fail to suppress either PKR or GCN2. A caveat in these results is that the amino-terminal boundary is not exactly the same for the various constructs, therefore we will express several different PKR and HRI constructs varying the amino-terminus of the protein to see if any of these truncated kinase domains can inhibit the activity of a wild-type eIF-2 α kinase (GCN2 or PKR).

4. Substrate specificity of the eIF-2 α kinases. This project is close to completion. A quantitative measure of the ability of the various combinations of kinases and substrates to activate GCN4 expression will be obtained by measuring the expression from a GCN4-lacZ allele. In addition, to even more convincingly show that the eIF-2 α -S51Y protein is phosphorylated on a tyrosine residue, extracts from various cells will be treated with phosphotyrosine-specific phosphatases prior to analysis using the isoelectric focusing gels. The phosphotyrosine phosphatase should specifically affect eIF-2 α -S51Y but not the wild-type protein. Mutant alleles of the GCN2 or PKR kinase that more efficiently phosphorylate eIF-2 α -S51Y could be identified to increase our understanding of the structural differences that distinguish the tyrosine-specific from the serine/threonine class of protein kinases, however, these projects are not planned for the immediate future.

5. Molecular analysis of yeast FUN12. The recent report of the DNA sequence of the entire yeast chromosome I identified a novel gene termed FUN12 that shares amino acid sequence similarity to bacterial IF-2 proteins. This protein is not the yeast mitochondrial IF-2 because this latter gene has already been identified elsewhere in the genome. IF-2 in bacteria functions to deliver the Met-tRNA_i^{Met} to the ribosome, the same role performed by eIF-2 in eukaryotes. Since yeast already contains eIF-2, the function of FUN12 is unknown. The translational regulation of GCN4 is exquisitely sensitive to translation initiation levels and especially the levels of Met-tRNA_i^{Met}, so the regulation of GCN4 expression is an ideal system to study FUN12.

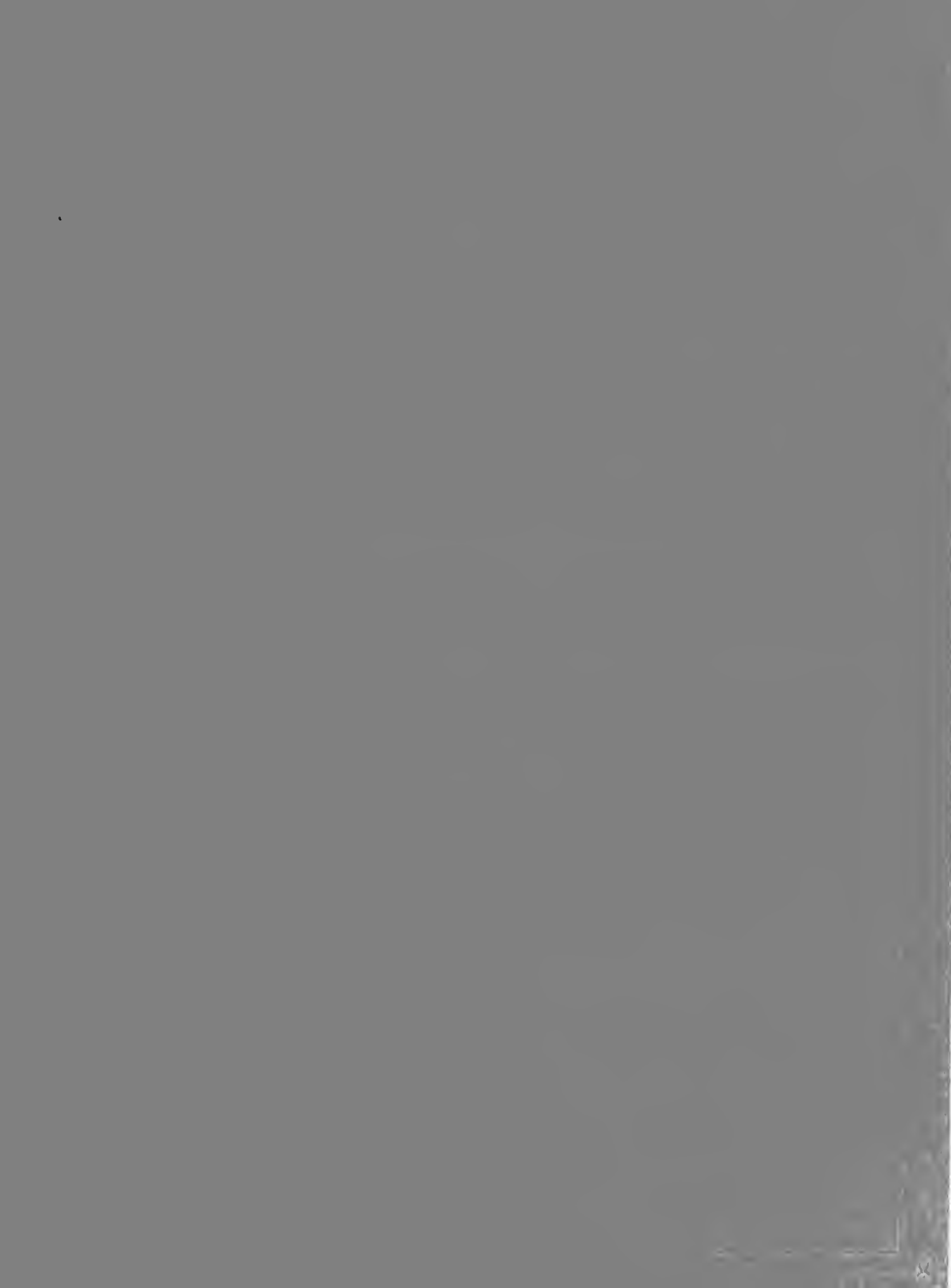
Preliminary results have revealed that overexpression of FUN12 will suppress the slow-growth phenotype of at least one of the GCN2^c kinases. This early result suggests that FUN12 functions in cytoplasmic translation initiation in yeast. To further explore the role of FUN12 in translation initiation, the FUN12 gene will be randomly mutated and the mutants will be screened to identify alleles that alter GCN4 translational regulation. We will screen for both alleles that block induction of GCN4 expression under amino acid starvation conditions as well as alleles that lead to high unregulated levels of GCN4 even in non-starvation conditions. The identification of either class of mutation will confirm that FUN12 functions in translation initiation. During this screen we will also identify any temperature sensitive (ts⁻) alleles of FUN12. The biochemical analysis of FUN12 ts⁻ mutants should help define the step of translation at which FUN12 acts. If the mutant screens fail to identify any interesting FUN12 alleles, the suppression of the GCN2^c kinases can be used in an analysis of the structural domains of the FUN12 protein. In addition to random mutants, site-directed mutations will be introduced into the FUN12 GTP-binding motif. Mutations in this region of bacterial IF-2 have been identified that lead to a slow-growth phenotype, we will examine the effects of the corresponding mutations in yeast FUN12. The identification of a protein in yeast showing homology to bacterial IF-2 proteins is quite surprising, through molecular genetic analysis of this gene in yeast we hope to identify the role for this protein in eukaryotic translation. These results take on added significance with our recent identification of an expressed sequence tag (EST) from mouse that bears striking homology to FUN12, raising the possibility that this protein is conserved in humans as well.

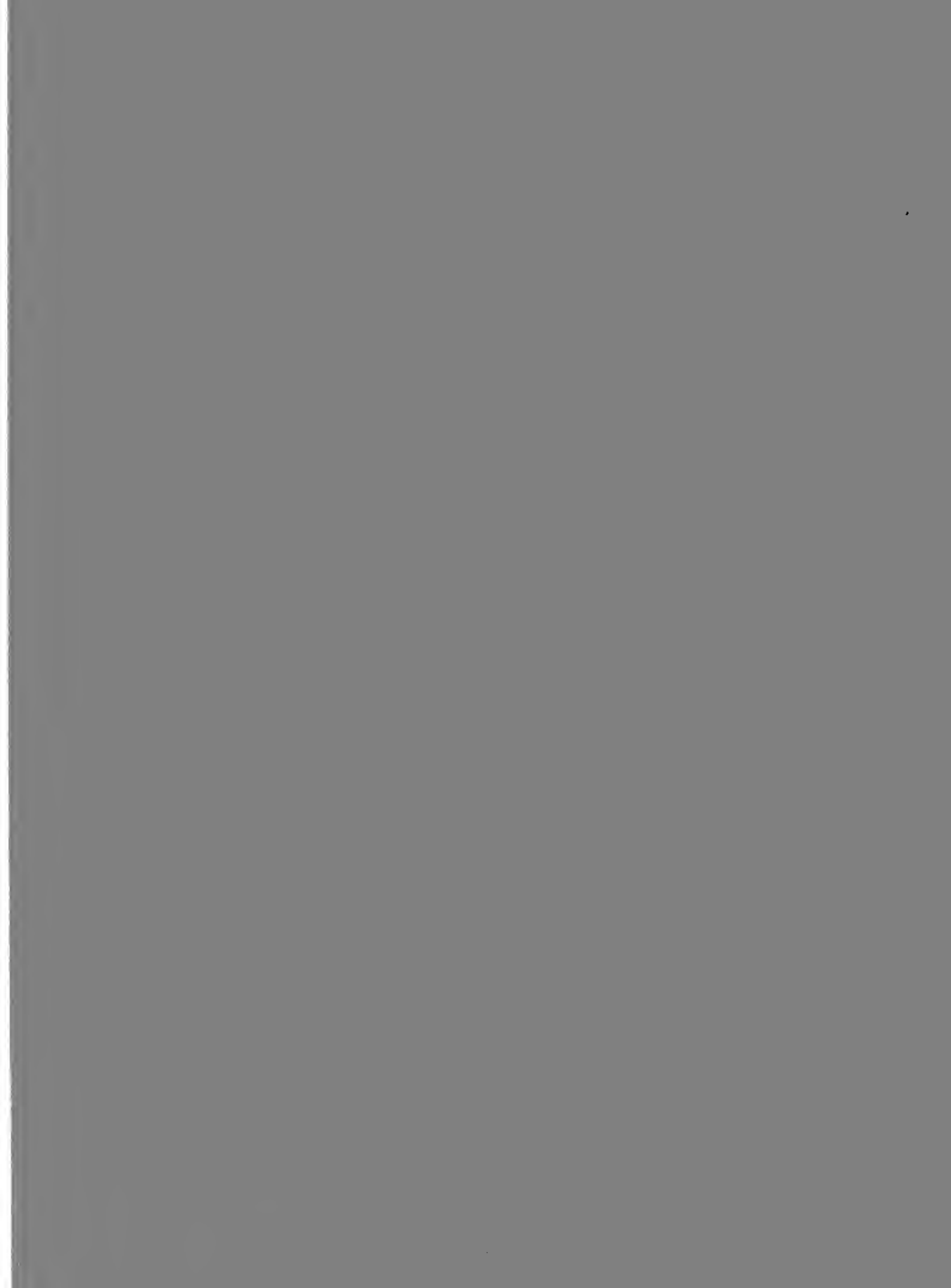
Significance to Biomedical Research and the Program of the Institute.

The eIF-2 α kinases play a key role in the antiviral and stress responses of mammalian as well as yeast cells. The ability of cells to alter gene expression under stress conditions is essential to the survival of the organism. In addition, the genetic characterization of substrate recognition by the eIF-2 α kinases should provide insights into the mechanism of substrate recognition by other protein kinases. Finally, increasing our understanding of viral inhibitors of the eIF-2 α kinases will not only provide insights into the molecular mechanisms regulating these kinases, but may also identify ways to defeat these viral defense mechanisms.

Publications:

Dever TE, Yang W, Astrom S, Bystrom AS, Hinnebusch AG. Modulation of tRNA_i^{Met}, eIF-2 and eIF-2B expression shows that GCN4 translation is inversely coupled to the level of the eIF-2•GTP•Met-tRNA_i^{Met} ternary complexes, Mol. Cell. Biol. (In Press).





DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD-01800-05 LMGD

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

The Genetic Basis of Mammalian Kidney Development

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute)

Gregory R. Dressler, Ph.D., LMGD, Head, Unit of Molecular Embryology

Uwe W. Rothenpieler, M.D., Visiting Associate, LMGD/NICHD

Larry T. Patterson, M.D., IRTA Fellow, LMGD/NICHD

Gabriella Ryan, Ph.D., IRTA Fellow, LMGD/NICHD

Dawn E. Phelps, M.S., Pre-IRTA Fellow, LMGD/NICHD

COOPERATING UNITS (if any)

Urologic Oncology Section, Surgery Branch, NCI (J. Gnarra); The Wistar Institute, Philadelphia, PA (F.J. Rauscher)

LAB/BRANCH

Laboratory of Mammalian Genes and Development

SECTION

Unit of Molecular Embryology

INSTITUTE AND LOCATION

NICHD, NIH, 9000 Rockville, Pike, Bethesda, MD 20892

TOTAL STAFF YEARS:

4 years

PROFESSIONAL:

3 years

OTHER:

1

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither

☐ (a1) Minors

☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Project Terminated.

LMGD-Fy95

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01-HD-01801-06-LMGD

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

The Molecular Analysis of Murine Development

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

K.A. Mahon, Ph.D., Expert, LMGD, NICHD

D.E. Feltner, M.D., Guest Worker

T. Hatta, M.D., Guest Worker

E. Hermes, Visiting Associate

L. Kos, Ph.D., Visiting Fellow

M.A. O'Reilly, Ph.D., NRC Fellow

COOPERATING UNITS (if any)

CBER, FDA (M. Jamrich), LP, NCI (S. Mackem), LMG, NICHD (T. Sargent), U. of Michigan (S. Camper)

LAB/BRANCH

Laboratory of Mammalian Genes and Development

SECTION

Unit on Developmental Gene Regulation

INSTITUTE AND LOCATION

National Institute of Child Health and Human Development, NIH, Bethesda, MD 20892

TOTAL STAFF YEARS:

4.75

PROFESSIONAL:

4.75

OTHER:

0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither

☐ (a1) Minors

☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Homeobox genes occupy key positions in the regulatory gene hierarchy responsible for establishing the embryonic body plan in *Drosophila*, and are thought to play analogous roles in the development of higher vertebrates. The homeobox gene *Rpx* is expressed during gastrulation in the prospective anterior neural plate and, at later stages, in Rathke's pouch, the primordium of the anterior and intermediate lobes of the pituitary. Extinction of *Rpx* expression coincides with the differentiation of pituitary-specific cell types. Using *lacZ* reporter genes, it has been shown that proper spatial and temporal expression in the anterior neural plate can be recapitulated in transgenic mice by as little as 600 bp of upstream sequence and a conserved 180 bp element in the first intron. A region at the 3' end of the gene is required for expression in Rathke's pouch. In addition, an element has been uncovered that directs transgene expression to a region of the hypothalamus and incipient posterior lobe that is in direct contact with Rathke's pouch. *In vitro* tissue recombination experiments have established that this expression is "induced" by contact with the pouch. It is proposed that this element may be present in other genes that normally respond to signals emanating from the pouch during the determination of the hypothalamic-pituitary axis. Molecular dissection of this element is in progress.

The role *Rpx* plays in pituitary development has been investigated in gain-of-function mouse models. Persistent expression in the pituitary has been achieved in transgenic mice by directing expression of *Rpx* with pituitary-specific promoters. Transgenic mice bearing α -glycoprotein subunit promoter-*Rpx* fusion genes express *Rpx* in the pituitary after the endogenous gene normally turns off. These mice are dwarfed and have underdeveloped reproductive tracts. Their pituitaries are very small and are deficient in gonadotroph, thyrotroph, lactotroph, and somatotroph cell lineages. These results indicate that the *Rpx* gene product must be downregulated during development for proper differentiation of the pituitary. It has been found that the expression of *Rpx* is not properly silenced during pituitary development in the Ames (*df*) dwarf mutant, which is phenotypically similar to the α -GSU-*Rpx* transgenic mice. These data indicate that *Rpx* mis-expression may be causatively involved in the genesis of certain types of pituitary dysfunction in mouse and man.

Project Description: The Role of Homeobox Genes in Early Vertebrate Development

Objectives:

The goal of this project is to identify genes that regulate early development in the mouse and to investigate their function.

Methods Employed:

Techniques utilized include the isolation and characterization of DNA and RNA, cDNA synthesis, molecular cloning in prokaryotic vectors, polymerase chain reaction (PCR) amplification, gene transfer technology, zygote microinjection, mouse surgery and dissection, tissue explant recombination and culture, the production of chimeras by blastocyst microinjection and embryo transfer, and *in situ* hybridization and immunostaining of tissue sections and whole mount preparations of embryos.

Major Findings and Proposed Course:

This group is involved in the molecular characterization of mammalian pattern formation, focusing on the early period of postimplantation development, from gastrulation through early neurogenesis. Research efforts have concentrated on divergent members of the phylogenetically conserved homeobox gene family, widely believed to be important developmental regulators in a variety of animal species. These genes can be sub-divided into several classes by virtue of sequence similarities in the homeobox DNA binding domain. The largest class of mammalian homeobox genes, the *Hox* genes, are analogous to the homeotic genes of the *Drosophila Antennapedia* and *Bithorax* clusters, and display region-specific expression in the developing central nervous system (CNS) and axial skeleton of the mouse embryo. Based on the extreme phylogenetic conservation of structure and expression pattern, as well as the phenotypes of gain-of-function and loss-of-function mouse models, it is evident that the *Hox* genes provide positional information along the embryonic axis. *Hox* gene activity is not required for the patterning of the head and the most rostral portion of the CNS, the prosencephalon or developing forebrain because *Hox* gene expression does not extent rostral to the

hindbrain. The cephalic portion of the CNS clearly develops differently than the rest of the CNS (for example, without the inductive influence of the notochord), indicating that there may be a group of genes uniquely required for its regionalization. In addition, transcripts from individual *Hox* genes are not detectable until after gastrulation is well underway, and accordingly do not appear to play a role in early determinative decisions, such as the establishment of polarity, embryonic axes, mesoderm and neural induction, or the initiation of morphogenesis. The functional characterization of new homeobox genes that may play a role in these processes has been the major focus of this group.

Several candidate homeobox genes have been isolated in the laboratory, including members of the *Distal-less* (*Dlx*) gene family and *Rpx*, which are unique in that they are restricted in their expression to the cephalic portion of the embryo. Another homeobox gene, *Nkx-3.1*, is also being investigated because it appears to be involved in the early determination of cell lineages in the somite. Current research centers on analysis of the regulation and function of these genes during embryogenesis.

The murine *Distal-less* homeobox gene family. Members of the murine homeobox gene family *Dlx*, homologous to the *Distal-less* gene of *Drosophila*, have been previously cloned by the group. The *Dlx* genes are expressed during embryonic development in the head of the developing embryo, particularly in the ventral forebrain and in the branchial arches, neural crest derived structures that give rise to the cartilage and skeleton of the face. Focusing on the *Dlx-2* and *Dlx-3* genes, the group has demonstrated that these genes are differentially expressed both temporally and spatially in the branchial arches and in subsets of their derivative, suggestive of a role in the patterning or differentiation of these structures.

Current work on the *Dlx-3* gene has centered on its regulation and its function during murine development. In collaboration with T. Sargent (LMGD, NICHD), it has been shown that approximately 1.5 kb of upstream sequence from the *Xenopus* *Dlx-3* homolog was sufficient to direct correct epidermal expression of a *lacZ* reporter gene in transgenic mice. The frog *Dlx* regulatory region was able to convey appropriate expression in the mammary gland and whisker follicles, structures that express the mouse gene but that do not exist in *Xenopus*. Thus, the regulatory regions of *Dlx* homeobox genes from very disparate vertebrate species are functionally conserved. The murine *Dlx-3* gene is currently being ablated using embryonic stem cell technology to ascertain the developmental function of this

gene.

***Rpx*, a homeobox gene expressed in the anterior neural plate and Rathke's pouch.** Another gene isolated in the laboratory, *Rpx* (for Rathke's pouch homeobox), is distinctive in that it is expressed early in postimplantation development in a very anterior-restricted domain. The *Rpx* gene is expressed during gastrulation (beginning at 6.5 days) in the prospective cephalic neural plate of the embryo. As development proceeds, transcripts become progressively restricted to Rathke's pouch, the primordium of the anterior and intermediate lobes of the pituitary. *Rpx* expression in Rathke's pouch is downregulated between 13.5 and 15 days of embryogenesis, coincident with the differentiation of pituitary specific cell types. This striking pattern of expression suggests a role in the determination and differentiation of the pituitary. However, since expression is detected several days prior to the development of Rathke's pouch, and initially extends over a much broader area than that destined to become the anterior pituitary, it seems likely that *Rpx* may also be involved in defining the anterior region of the embryo.

The group is examining both the regulation and function of the *Rpx* gene during development. *Rpx* gene regulation has been studied in transgenic mice by introducing reporter constructs that contain *lacZ* fused in frame to potential regulatory sequences from the *Rpx* gene. Histochemical staining of transgenic embryos has indicated that the early expression pattern in the prechordal plate and cephalic neural plate (7-9 days) can be achieved with as little as 600 base pairs of upstream sequence and the first intron. Analysis of the sequence of these regions revealed the presence of several known transcription factor binding sites and at least two unique motifs that are also present in the frog *Rpx* homolog (studied by M. Jamrich). Deletion of one of these conserved motifs, a 180 bp GAGA repeat in the first intron, totally eliminated expression, indicating its importance for *Rpx* regulation. Expression in Rathke's pouch requires an additional element between 3-5 kb downstream of the gene. Thus, at the level of gene regulation, early stage expression of *Rpx* is distinct from late stage expression.

Interestingly, at later stages of development, expression of transgenes that lacked the Rathke's pouch-specific region was detected in the neuroectoderm of the hypothalamus that lies in direct contact with the pouch, a region where the endogenous gene is not normally expressed. This curious pattern of transgene expression was intriguing because of its proximity to Rathke's pouch, and suggested that expression was elicited by cell-cell contact. Tissue recombination experiments

have verified that *lacZ* expression in ectopic sites of the neuroectoderm can be induced in co-cultures with isolated Rathke's pouches from nontransgenic animals. Deletion analysis in transgenic mice has indicated that a 200 bp region (-600 to -400) is required for this "inducible" expression. It has been postulated that reciprocal inductive interactions between the pouch and the floor of the hypothalamus are required for the proper determination and differentiation of the hypothalamic-pituitary axis. The "inducible" element unmasked in these transgene studies may be present in other genes that normally respond to signals emanating from the pouch, and conceivably may be used to isolate such genes.

The inducible element is currently being scrutinized by deletion and gel-shift analysis to determine more precisely which sequences are responsible for the inductive response. To test whether this region behaves like a bona fide enhancer, the 200 bp region was placed upstream of a heterologous promoter fused to *lacZ* and injected into mouse embryos. Surprisingly, a significant proportion of the transgenic progeny had malformations of the axis. The most common phenotype was microcephaly, which results presumably from a reduction of prechordal plate material. Two embryos had duplications of the axis, a relatively common malformation in the frog, but extremely rare in the mouse. Since there are many lines of mice that express *lacZ* ubiquitously during development without malformation, the working hypothesis is that the *Rpx* element, when removed from its normal context, is titrating out a limiting transcription factor expressed in the node, notochord, or prechordal plate. To test this model, the element has been multimerized and is being injected without *lacZ* sequences into both mouse and frog embryos. Embryos will be scored for axial malformations and the copy number of the transgene determined. This approach, coupled with the biochemical analysis of the sequence, should provide insight into the molecular mechanisms of axis formation in vertebrates.

The developmental function of *Rpx* has been tested in gain-of-function mouse models. The promoter for the α subunit of the glycoprotein hormones (α GSU) has been used to express the *Rpx* gene in transgenic mice in order to gauge the effects of persistent expression in the pituitary. The α GSU gene is normally expressed in Rathke's pouch in most or all of the progenitor cells of the pituitary at day 9 and continues to be expressed in the pituitary in gonadotrophs and thyrotrophs after *Rpx* turns off. Transgenic mice carrying the α GSU-*Rpx* fusion gene are dwarfed and hypogonadal, and have very small pituitary glands. Immunostaining has revealed that virtually all of the cell types of the anterior pituitary (thyrotrophs,

gonadotrophs, somatotrophs, and lactotrophs) are severely depleted, with the exception of corticotrophs. This result indicates that *Rpx* must be downregulated for proper differentiation of the pituitary. Persistent *Rpx* expression appears to inhibit differentiation of most of the cell types derived from α GSU expressing progenitor cells, suggesting that it perturbs the derivation and proliferation of cell lineages of the pituitary.

There are a number of mutations in the mouse that phenotypically resemble these transgenic mice. Most notable among these are the *Snell* (*dw*) and the *Ames* (*df*) dwarf mice. These mutants are deficient in three pituitary cell lineages, thyrotrophs, somatotrophs, and lactotrophs, and also have hypoplastic anterior pituitary glands. The *Snell* dwarf is known to result from a mutation in the homeodomain of the pituitary-specific transcription factor *Pit-1*. The molecular lesion in the phenotypically indistinguishable *Ames* dwarf is unknown. In collaboration with Dr. S. Camper (University of Michigan), the group has investigated whether the expression pattern of the *Rpx* gene is altered in these mutants.

The rationale for this investigation is based on the following observations. There are several binding sites for *Pit-1* present in region upstream of the *Rpx* gene. *Pit-1* is known to regulate the transcription of the growth hormone and prolactin genes, but it performs a less well understood role in the proliferation and differentiation of several pituitary cell lineages. Comparative *in situ* hybridization experiments have shown that expression of *Rpx* is extinguished at the same time that *Pit-1* is activated, consistent with the hypothesis that *Pit-1* negatively regulates *Rpx* expression in the developing pituitary. However, it was found that *Rpx* is regulated appropriately in the *dw* mutant, indicating that *Pit-1* is not uniquely required for the downregulation of *Rpx*. Interestingly, *Rpx* is not downregulated in the *Ames df* mutant. By analogy to the transgenic overexpression studies, persistent expression of *Rpx* may be causatively involved in the genesis of the *df* phenotype. This result for the first time establishes that the *Ames* gene product is required early in pituitary development, before the requirement for *Pit-1*, and is necessary for silencing *Rpx* expression. Failure to downregulate *Rpx* in development could potentially explain certain idiopathic hypopituitarism in man, and the group plans to investigate this possibility in collaboration with Drs. George Chrousos and L. Neiman (DEB).

***Nkx-3.1*, a homeobox gene dynamically expressed during somitogenesis.** In contrast to the *Dlx* and *Rpx* genes that are expressed in the anterior or cephalic

region of the embryo, the *Nkx-3.1* homeobox gene is dynamically expressed at the caudal end of the embryo in the youngest, most newly formed somites. The localization of *Nkx-3.1* transcripts within each somite varies depending on its position along the rostrocaudal axis. Expression is initially seen in the more medial regions of the somite that are in contact with the notochord and neural tube. In slightly older, more rostral somites, the gene is transiently expressed in the nascent myotomal cells at the rostral and caudal borders of each somite. There is a great deal of evidence from experimental embryology that signals from the notochord are required for proper differentiation of the sclerotome and myotome cell lineages. This is the only known gene that is expressed in a manner that suggests it is directly responding to these signals. To test this, mice homozygous for the *Danforth's short tail* mutation (*Sd*) were hybridized with *Nkx-3.1* probes. These mice are characterized by degeneration of the notochord in the caudal region, but do have morphologically identifiable somites. It was found that the *Nkx-3.1* gene was not expressed in regions that lacked a notochord, indicating that this structure is required for proper expression of this gene in the somite.

The role of the axial structures for *Nkx-3.1* expression was examined more directly using a mouse somite explant system developed by the group. *Nkx-3.1* was only expressed in somite explants that were cultured in combination with the notochord and/or neural tube. Recently, it has been shown that a short-range signalling molecule, *sonic hedgehog*, is expressed in the notochord and is capable of eliciting floor-plate formation in the developing spinal cord, a function known to be imparted by the notochord. The group has found that *sonic hedgehog* is sufficient to elicit expression of *Nkx-3.1* in isolated somites in the absence of the notochord. The dynamic expression pattern of *Nkx-3.1* and its requirement for signals from the notochord suggest that it plays an important instructional role in translating these signals within the developing somite to bring about the derivation of somitic cell lineages.

Significance to Biomedical Research and the Program of the Institute:

This work will focus on the mechanisms governing both normal and abnormal mammalian development and may ultimately lead to a better understanding of human birth defects and endocrine abnormalities.

Publications:

Hermesz E, Mahon KA. Rpx: A novel anterior-restricted homeobox gene progressively activated in the prechordal plate, anterior neural plate, and Rathke's pouch of the mouse embryo. Development (in press)

Morasso MI, Mahon KA, Sargent TD. A Xenopus Distal-less gene in transgenic mice: Conserved regulation in distal limb epidermis and other sites of epithelial-mesenchymal interaction. Proc Nat Acad Sci USA 1995;9:3968-72.

Morgenbesser SD, Schreiber-Agus N, Bidder M, Mahon KA, Horner J, DePinho RA. Contrasting roles for c-Myc and L-Myc in the regulation of cellular growth and differentiation in vivo. EMBO J 1995;14: 743-56.

Ransom M, Tickle C, Mahon KA, Mackem S. A new chick gene related to the organizer associated homeobox gene Xnot, may regulate positional identity along the limb proximodistal axis. Mech Devel 1995;51:17-30.

Robinson GW, Mahon KA. Differential and overlapping expression patterns of Dlx-2 and Dlx-3 suggest distinct roles for Distal-less homeobox genes in craniofacial development. Mech Devel 1994;48:199-215.

Scherer SW, Robinson GW, Mahon KA, Heng HHQ, Poorkaj P, Evans JP, Tsui L-C. Assignment of the human homolog of mouse Dlx-3 to chromosome 17q21-q22 by analysis of somatic cell hybrids and fluorescent in situ hybridization. Mammalian Genome 1995;6:310-11.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01-HD-01803-02-LMGD

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Genetic analysis of Thymocyte Development

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

Paul E. Love, MD, PhD; Investigator (PHS), LMGD, NICHD

Connie L. Sommers, PhD; IRTA Fellow, LMGD, NICHD

Hala Azzam, Ph.D.; IRTA Fellow, LMGD, NICHD

Kun Huang, MS; Biologist, LMGD, NICHD

Howard Shen, Pre-IRTA Fellow, LMGD, NICHD

Daniel A. Charlick, Pre-IRTA Fellow, LMGD, NICHD

COOPERATING UNITS (if any)

CBER, FDA (E.W. Shores) STR, LMGD, NICHD (H. Westphal, A. Grinberg, E.J. Lee)

LMAI, NIAID (V. Flamand, J.-P. Kinet) LMI, NHLBI (X. Cao, W. J. Leonard)

LAB/BRANCH

Laboratory of Mammalian Genes and Development

SECTION

Unit on Cellular and Developmental Biology

INSTITUTE AND LOCATION

NICHD, NIH 9000 Rockville Pile, Bethesda, MD 20892

TOTAL STAFF YEARS:

5.25

PROFESSIONAL:

3.0

OTHER:

2.25

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither

☐ (a1) Minors

☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Research is directed at understanding the cellular and genetic events that control normal T cell development. Transgenic and gene-targeting methods are used to analyze the function of known genes and various techniques (eg., RT-PCR) are being employed to identify novel genes that participate in thymocyte development. Current studies are focused on: (I) The role of the T cell antigen receptor (TCR) in thymocyte maturation. In mature T cells, the TCR transduces signals important for T cell activation and cell mediated immunity and in immature T cells the TCR directs thymocyte development and thymic selection. The TCR is composed of multiple signal transducing subunits (the CD3 chains and one or more members of the ζ -family of proteins; ζ , η and Fc ϵ RI γ) that couple the TCR to intracellular signal transduction pathways. To address whether these signal transducing subunits perform distinct or analogous functions in development we have: a) examined the function of each of the individual subunits (ζ , CD3- ϵ , Fc ϵ RI γ , η) by overexpression in transgenic mice; b) examined the role of ζ -chain in T cell ontogeny by generating ζ/η -deficient and subsequently ζ/η /Fc ϵ RI γ deficient mice, and c) examined the role of the 3 individual ζ -chain signaling motifs (ITAMs) in thymocyte development by transgene reconstitution of ζ -deficient mice. These studies have revealed that expression of ζ chain (or a related protein) is required for TCR surface expression but that ζ chain signals are not specifically required for T cell development. Thus the CD3 subunits are sufficient to transduce all of the signals necessary for thymocyte maturation. Moreover, the ζ and CD3 subunits contribute to the overall signaling potential of the TCR and together determine the specificity of the T cell repertoire generated in the thymus. (II) The role of other signal transducing proteins in T cell development is being examined by the generation of transgenic mice. These include CD5 (which also contains an ITAM), and the protein tyrosine kinase (PTK) ZAP-70, which is known to interact with ζ chain. (III) Novel genes that have potential functions in thymocyte development or T cell activation are being identified. A new protein tyrosine kinase gene, *txk* has been cloned and biochemical and transgenic approaches are being employed to analyze its function. A similar approach is also being used to identify genes involved in early thymus/thymocyte development.

Project Description: Genetic analysis of T Cell Development

Objectives: To understand the cellular and molecular events that control T cell development and formation of the mature T cell repertoire.

Methods Employed: Transgenic and gene targeting methods are used to create overexpression, dominant-negative, and loss-of-function mutants to investigate the role of specific genes in T cell ontogeny. Molecular and genetic techniques (including generation and screening of cDNA libraries, PCR and cloning) are also being used to identify novel genes that function in thymocyte development. Biochemical studies include: cell surface radioiodination / immunoprecipitation / western blotting / 1- and 2-dimensional SDS-PAGE, NEPHGE. Cell culture techniques include: fetal thymic organ culture and transfection of established cell lines. Flow cytometric analysis of thymocytes and peripheral T cells. Immunohistochemistry and *In situ* hybridization.

Background

T cells develop throughout the life of an individual. Progenitor cells originate in the fetal liver or adult bone marrow and then migrate to thymus where they mature and reach their full potential. The most immature thymocytes are designated double negative (DN) as they do not express the CD4 and CD8 coreceptors. Rearrangement of the T cell antigen receptor (TCR) β locus is initiated in DN thymocytes and triggers a seemingly irreversible lineage commitment leading to the formation of cells that express both CD4 and CD8 (double positive, DP). Rearrangement of TCR α occurs in DP thymocytes and these cells then express surface $\alpha\beta$ TCR. DP thymocytes then undergo positive and negative selection, a process that directly depends upon the specificity of their TCR and that leads to the survival of competent "self-educated" cells, deletion of autoreactive cells and death of nonresponsive cells. The ultimate product of the thymocyte maturational pathway is formation of mature functional cells that express CD4 or CD8 but not both (i.e., single positive, SP) and high levels of surface $\alpha\beta$ TCR.

Differentiation of precursor thymocytes into mature, functional T cells is controlled in part by signals delivered through the TCR. In particular, TCR mediated signals regulate positive and negative selection in the thymus, and are required for the differentiation of immature thymocytes into mature functional T cells. The TCR complex that is expressed on DP and SP thymocytes and most T cells is composed of six different subunits that function either in antigen recognition or signal transduction. The clonotypic TCR α/β chains are responsible for ligand specificity and are generated by productive rearrangement of germline gene segments, a process mediated by the products of the Rag-1 and Rag-2 genes. The α/β heterodimer lacks inherent signaling activity but associates noncovalently with multiple signal transducing subunits: the CD3- γ , - δ and - ϵ components, and a ζ chain homodimer. ζ belongs to a family of proteins (ζ , η and Fc ϵ R1 γ) that are structurally and functionally related. The CD3 and ζ family proteins contain semi-conserved sequences within their cytoplasmic domains, termed *Immunoreceptor Tyrosine-based Activation Motifs* (ITAMs) that couple the TCR to intracellular signal transduction pathways by their ability to interact with protein

tyrosine kinases. The CD3 and Fc ϵ R1 γ chains each contain a single ITAM, η contains two ITAMs and ζ contains three ITAMs.

The rationale for multiple TCR signal transducing subunits has been enigmatic. In particular, it is unclear whether the ζ and CD3 molecules perform identical or distinct signaling functions, or if the relative contribution of these molecules varies with the developmental stage of the T cell. The multisubunit nature of the TCR suggests the potential for variability in the signaling response depending upon its configuration. Indeed, a precursor form of the TCR (the pre-TCR) is expressed on DN thymocytes. The pre-TCR, which is expressed prior to rearrangement of TCR α , consists of the β chain paired with a pre-T α chain (pre-T α) and the CD3 chains; ζ chain either does not associate with the pre-TCR or is only loosely associated with this complex. It has also been shown that the TCRs expressed on a subpopulation of mature T cells ($\gamma\delta$ T cells) contain both ζ and Fc ϵ R1 γ chains raising the possibility that inclusion of different ζ -family molecules as part of the TCR complex could result in the transduction of distinctive signals following TCR engagement.

We have begun to investigate the role of the TCR signals in T cell development and T cell function by analyzing the *in vivo* function of the various TCR signal transducing subunits. Our approach has been to use transgenic and gene targeting methods to create overexpression, dominant-negative, and loss-of-function mutants.

Major Findings:

1. Role of TCR signaling molecules in T cell development (in collaboration with EW Shores)

a. Overexpression of ζ -chain in transgenic mice. Transgenic mice that overexpress each of the ζ -family proteins (ζ , η or Fc ϵ R1 γ) were generated in our facility. Analysis of ζ transgenic (ζ Tg) mice yielded the unexpected finding that thymocyte development was arrested at the early DN precursor stage. Further investigation revealed that overexpression of ζ -chain caused premature termination of Rag-1 and Rag-2 expression in early thymocytes, preventing productive rearrangement of the TCR α and TCR β genes. Interestingly, overexpression of truncated ζ chains or other ζ -family proteins (η , Fc ϵ R1 γ) in early thymocytes did not result in downregulation of Rag expression, indicating that this effect was specific to ζ . Down-regulation of Rag-1 and Rag-2 in ζ Tg mice appears to be the result of ζ -mediated signals; coinciding with the down-regulation of Rag transcripts, DN thymocytes from ζ transgenic mice express inappropriately high levels of the CD5, CD2 and CD69, antigens whose expression is known to increase upon T cell activation. Interestingly, the effects of ζ -chain overexpression (i.e., thymocyte activation and downregulation of Rag-1 and Rag-2) are similar to those normally observed in $\alpha\beta$ TCR⁺ DP thymocytes as they undergo thymic selection and then mature into SP thymocytes. Signals transduced by the $\alpha\beta$ TCRs on DP thymocytes regulate expression of Rag-1 and Rag-2 as direct engagement of $\alpha\beta$ TCR complexes on DP thymocytes results in down-regulation of Rag-1 and Rag-2. Termination of Rag expression is thought to be a critical step in late thymocyte development as it effectively results in allelic exclusion of TCR α , thus fixing the specificity of the $\alpha\beta$ TCR.

Immature DN thymocytes express a precursor form of the TCR pre-TCR complex (pre-TCR) which consists of TCR β , a pre-T α chain and the CD3 chains. The pre-TCR either does not associate with ζ or only weakly associates with ζ . If ζ -chain were part of the pre-TCR complex (as we predict may occur in transgenic mice overexpressing ζ), signals mediated by this complex could prematurely down-regulate Rag expression preventing rearrangement of TCR α . These results suggest that a ζ -less pre-TCR complex may be necessary for the transduction of signals which support early thymocyte development, but which do not prematurely terminate Rag expression. Signaling through the pre-TCR controls differentiation of DN thymocytes into DP thymocytes and also controls allelic exclusion of TCR β . Therefore, the configurations of the pre-TCR and the $\alpha\beta$ TCR appear to be particularly well adapted to transduce signals that control allelic exclusion of the TCR β and TCR α genes, respectively. Our results indicate that inclusion of ζ chain as a signaling component of these surface complexes may be developmentally regulated so that the appropriate signals are delivered to thymocytes at specific stages of development.

b. Generation of $\zeta^{-/-}$ mice. Examination of mice in which the genes encoding ζ -chain have been inactivated by homologous recombination ($\zeta^{-/-}$) revealed the importance of this protein in normal T cell development. Thymi from $\zeta^{-/-}$ mice are markedly reduced in size and cellularity and the phenotype of thymocytes reveals that in the absence of ζ -chain, T cell development is inhibited. In addition, thymocytes and T cells from $\zeta^{-/-}$ mice express barely detectable levels of TCR on their surface. SP thymocytes are almost undetectable in $\zeta^{-/-}$ mice, however, a small number of these cells can be identified. Yet, despite the almost complete lack of SP thymocytes in the thymus, the periphery of these animals contain low, but readily detectable number of SP T cells that seem to accumulate with age. The apparently normal ratio of CD4 $^{+}$ to CD8 $^{+}$ T cells supports the idea that these cells are thymically derived. The forces driving the development of these SP cells remain unclear since they express extremely low levels of surface TCR. One possibility is that SP cells develop through a TCR independent pathway, perhaps involving signals through molecules such as Thy1, Ly6, CD2 or CD5. However, it may not be necessary to invoke alternative signaling pathways as we were able to detect low levels of CD3 ϵ and TCR β on the surface of T cells from $\zeta^{-/-}$ mice by flow cytometry. Moreover, we have found that DP thymocytes from $\zeta^{-/-}$ mice express low levels of signaling-competent TCR complexes. Engagement of the TCRs on DP thymocytes from $\zeta^{-/-}$ mice with antibodies directed against the TCR β chain leads to signaling events associated with positive selection. Therefore, we favor the idea that the generation of SP thymocytes is a consequence of TCR mediated signals delivered through a "nonclassic" TCR composed of α , β and CD3 chains, but lacking ζ .

The implications of this finding are quite striking and unexpected as they suggest that engagement of a ζ -less TCR complex can deliver signals sufficient for the positive selection of at least some thymocytes. The relative inefficiency of this process (as reflected by the number of SP cells generated) could be reflective of the extremely low level of TCR expression. Interestingly, despite the ability of ζ -less TCRs to transmit signals associated with positive selection, peripheral T cells from $\zeta^{-/-}$ mice fail to proliferate in response to mitogens or TCR engagement. These findings could indicate an anergic state of these cells, as might be induced by TCR engagement in the absence of costimulatory requirements (e.g. IL-2). Another interpretation is that signals required for positive selection may be quantitatively or qualitatively different than those signals required for activation

of peripheral cells.

The fact that some DP thymocytes are generated in $\zeta^{-/-}$ mice is consistent with studies indicating that ζ -chain may not be part of the early pre-TCR complex. As these early complexes are thought to consist of TCR β /preT α associated with CD3 signaling components but not ζ , it is likely that the CD3 components alone can promote early thymocyte differentiation. In this regard, it has been shown by others that antibody cross-linking of CD3 components on DN thymocytes from $\zeta^{-/-}$, Rag-1 $^{-/-}$ double knockout mice induces their differentiation to the DP stage, demonstrating that ζ chain signals are not required for this step. It remains unclear why ζ -chain either does not associate or only weakly associates with this early complex as ζ is known to be expressed early in fetal development. One hypothesis is that replacement of pT α with TCR α is required for ζ -chain to become stably associated with the TCR complex. Since TCR α is rearranged and expressed at the DP stage, this would imply that ζ -chain does not become stably associated with the TCR complex until this stage of development, a speculation consistent with the phenotype of ζ -chain deficient mice. Interestingly, unlike $\zeta^{-/-}$ mice, TCR $\alpha^{-/-}$ mice and Rag $^{-/-}$ mice reconstituted with a TCR β transgene exhibit normal numbers of DP thymocytes. The reduction in DP thymocytes in $\zeta^{-/-}$ mice could reflect: 1) failure of DP thymocytes (or their immediate precursors) to expand in the absence ζ -chain signals, 2) inefficient transition of thymocytes from the DN stage to the DP stage, or 3) death of DP cells by "neglect" due to their inability to express sufficiently high levels of TCR. While the resolution of this paradox will require future experiments, it is also possible that ζ -chain might transduce signals independent of an association with the pre-TCR complex.

c. Genetic reconstitution of ζ -deficient mice. Because ζ is required for efficient surface expression of the $\alpha\beta$ TCR complex, it was impossible to discern whether the developmental defects in ζ -deficient mice were due specifically to the lack of ζ -mediated signaling or to the failure to target other TCR subunits to the cell surface. Thus, although the phenotype of ζ -deficient mice revealed that ζ chain plays an important role in thymocyte development, the relative significance of ζ -mediated signals, individual ζ -chain signaling motifs (ITAMs), or CD3 subunit-mediated signals could not be adequately addressed in $\zeta^{-/-}$ mice. One way to examine this fundamental question was to reconstitute $\zeta^{-/-}$ mice with ζ chains that either partially or completely lacked sequences required for signal transduction. To this end, we generated a set of transgenes that encoded either full length ζ chain or genetically engineered variants of ζ that lacked one or more ITAMs. The transgenes were placed under the control of identical regulatory sequences, and since they all retained the extracellular and transmembrane ζ chain domains, the encoded proteins were able to dimerize, assemble with other TCR components and fully restore surface TCR expression, thus enabling a direct analysis of the role of ζ -chain signals in T cell development.

Remarkably, all of the transgene encoded ζ -chain variants, most notably the variant which lacked all three ITAMs, alleviated the block in thymocyte development. Numbers of DP and SP thymocytes were markedly increased in these mice relative to $\zeta^{-/-}$ mice. Restoration of normal or near-normal numbers of SP thymocytes and peripheral T cells in all of the reconstituted lines indicated that even TCR complexes that contained signaling-deficient ζ chains (0 ITAMs) were capable of supporting positive selection of thymocytes. Moreover, these TCR complexes were competent to trigger cellular responses directly associated with positive selection, such as

downregulation of Rag-1 and Rag-2 mRNA. Thus, these experiments firmly established that none of the ζ -chain ITAMs are specifically required for thymocyte development or positive selection. In addition, the ability of the signaling-deficient ζ chain to rescue thymocyte development in $\zeta^{-/-}$ mice demonstrated that, providing TCR surface expression was restored to normal levels, signals transduced by the CD3 subunits are sufficient to mediate thymocyte maturation and selection.

Although these experiments did not identify a qualitative difference between individual ζ or CD3 subunit ITAMs, they did reveal more subtle quantitative effects. At similar levels, TCR complexes containing full length ζ -chains were more efficient at reconstituting T cell development (as judged by numbers of SP thymocytes and peripheral T cells) than complexes containing ζ -chains that lacked ITAMs. Thus, these data indicate that the multiple ζ -chain ITAMs may function primarily to amplify TCR signals. Signal amplification might be important for establishing the T cell activation threshold, but perhaps more importantly, for regulating thymocyte positive and negative selection. During development, the fate of thymocytes is ultimately dictated by their clonotypic TCR and the signals it transduces. However, it has been difficult to understand how signals delivered through the same TCR complex can differentially lead either to positive or negative selection. While a number of models have been proposed to explain these two different outcomes of TCR engagement, models based on the quantity or strength of the TCR signal seem most compatible with the results obtained with transgene-reconstituted $\zeta^{-/-}$ mice. Several factors that appear to quantitatively influence TCR signaling include 1) the nature of the ligand and how well it "fits" with the TCR, 2) the quantity of ligand thus influencing the number of TCRs engaged, and 3) the level of TCR or coreceptor (CD4 or CD8) surface expression. Whereas these models have primarily focussed on the nature of the TCR-ligand interaction, their prediction is that decreased or enhancing signaling can influence whether signals delivered through the TCR will result in positive or negative selection.

Preliminary data from our laboratory provide additional support for the quantitative model. We have bred into $\zeta^{-/-}$; Tg -reconstituted mice TCR $\alpha\beta$ transgenes. The TCR $\alpha\beta$ transgenes inhibit rearrangement of endogenous TCR α and TCR β genes, thereby producing a nearly uniform population of thymocytes with identical TCRs. The use of the TCR $\alpha\beta$ transgenic system is extremely powerful for assessing the effects of other factors on positive and negative selection as all thymocytes exhibit the same TCR specificity. Because the transgenic TCR recognizes an antigen (HY) expressed only in males, virtually all thymocytes are positively selected in female mice but negatively selected in male mice. Our results with this system indicate that the number of ζ -chain ITAMs influences the threshold for positive selection (although we expect that similar effects will also be observed on negative selection). These results additionally suggest that the CD3 and ζ chains contribute to the overall signaling potential of the TCR complex, and together influence the specificity of the T cell repertoire generated in the thymus.

d. Generation of Fc ϵ R1 γ transgenic mice. Analysis of transgenic mice overexpressing Fc ϵ R1 γ chain revealed a phenotype quite different than that of ζ overexpressing mice. Interestingly, these mice exhibit delayed development of $\alpha\beta$ T cells, $\gamma\delta$ T cells and TCR $^{-}$ Natural Killer cells. However, unlike ζ Tg mice, downregulation of Rag-1 and Rag-2 did not occur in γ Tg mice. DN thymocytes from Fc ϵ R1 γ -transgenic mice express unusually high levels of CD16 (Fc γ RIII) a receptor that is normally expressed on NK cells in association with Fc ϵ R1 γ . T cells and NK cells are thought to be

derived from a common (DN, TCR⁻ CD16⁺) precursor in the thymus. Thus, our results suggest that signals mediated by FcεRIγ may control very early steps in thymocyte development.

2. Identification of novel genes involved in T cell development.

We have recently initiated a search for novel genes that may be important for thymocyte maturation. Using a reverse-transcriptase PCR (RT-PCR)-based strategy that employs degenerate oligonucleotides specific for certain classes of proteins (including protein tyrosine kinases, protein tyrosine phosphatases and homeobox genes), RNA from early fetal (day 13) or adult thymus is being screened.

Txk. We have identified a novel nonreceptor protein tyrosine kinase gene expressed in early thymocytes (*txk*). Sequence analysis indicates that Txk contains SH2, SH3 and kinase catalytic domains and belongs to the Tec family of protein tyrosine kinases. Northern blot analysis revealed that Txk is expressed in thymocytes as early as fetal day 13.5 and its expression continues throughout development. *txk* transcripts are present in peripheral T cells and in mast cell lines but are not detectable in B cell, macrophage /monocyte cell lines or in non-hematopoietic fetal or adult tissues. In both thymocytes and T cells *txk* transcripts are downregulated after activation by TCR cross-linking but not by treatment with IL-2 which stimulates cell proliferation. Expression of Txk in COS cells confirms its cytoplasmic localization and kinase activity.

3. Generation mice lacking the common cytokine receptor γ chain. (in collaboration with W.J. Leonard, NIAID). Our laboratory has also collaborated on a project aimed at generating a mouse model for human severe X linked combined immunodeficiency (XSCID) by inactivation of the common cytokine receptor γ chain (γc). γc is signal transducing subunit shared by the IL-2, IL-4, IL-7, IL-9 and IL-15 receptors. The γc-deficient mice exhibit a wide range of interesting immunological abnormalities, some shared by humans with XSCID and some that are different. The differences are indicative of variations in lymphoid development between humans and mice, and potentially will allow new insights into the roles of γc-dependent cytokines in both species. Moreover, the mutant mice provide a valuable animal model of γc deficiency and therefore represent potential targets for reconstitution by gene therapeutic approaches.

Proposed course of projects.

1. Role of TCR signaling molecules in T cell development. We have also reconstituted the ζ^{-/-} mice with transgenes encoding the other ζ-family molecules (η and FcεRIγ) to determine if these signaling molecules confer distinctive signaling properties on the TCR. In these mice, virtually all of the TCRs expressed by thymocytes and peripheral T cells contain either η or FcεRIγ instead of ζ. Preliminary data indicate that thymocyte development proceeds in the absence of ζ, a result consistent with our early findings with ζ^{-/-} mice reconstituted with transgenes encoding truncated ζ chain variants. Ongoing studies involve functional characterization of the T cells generated by these

mice. In collaboration with J. Ravetch (Memorial Sloan-Kettering) we have also generated mice lacking all of the ζ -family proteins ($\zeta\eta\gamma^{-/-}$). Initial characterization of these mice indicates that they generate all of the peripheral $\alpha\beta$ T cell subsets, but these cells are TCR $^{-}$. Interestingly, these mice develop a form of inflammatory bowel disease (IBD) that resembles, in some respects, human IBD. In addition to further characterizing this disorder, studies on these mice will be directed at determining whether they reveal any functional overlap between ζ and Fc ϵ R1 γ by comparing their phenotype to both sets of single knockout mice.

2. Functional characterization of other signaling molecules. We have recently generated transgenic mice expressing wild-type and mutant (ITAM-less) forms of the CD3 ϵ chain to analyze the function of this TCR subunit in T cell development. This studies are viewed as complementing our ongoing characterization of TCR signaling subunits. We envision generating combinations of the $\zeta^{-/-}$, ζ Tg and CD ϵ Tg mice by matings to further characterize the function of this molecule. CD5 transgenic mice have also been generated. CD5 is a surface receptor expressed on thymocytes and mature T cells that is thought to participate in T cell activation and perhaps thymic selection. Like ζ , CD5 contains an ITAM within its cytoplasmic region. Although the function of this motif is unknown, it has been suggested that CD5 cooperates with the TCR during lymphocyte activation and that its signaling ability may depend upon its association with ζ . Utilizing our existing mice, we can test these hypotheses by matings similar to those described for CD3 ϵ .

3. Txk The restricted expression pattern of *txk* and its regulation by T cell activation make this gene an excellent candidate for involvement in signal transduction during T cell development. Ongoing characterization of *txk* includes the production of anti-peptide and GST-fusion protein antibodies. We also plan to transfect the *txk* gene into established T cell lines to search for associated proteins (substrates) and to further characterize its regulation by and possible role in T cell activation. Finally, we have begun to generate transgenic mice overexpressing wild-type and dominant negative (kinase dead) forms of Txk to determine its possible role in T cell development.

4. Genes controlling early thymocyte development. We have also utilized RT-PCR to identify potentially novel homeobox containing genes involved in early thymus/thymocyte development. We designed primers for several subfamilies of homeobox genes, but have concentrated first on the *Antennapedia* subfamily of homeobox genes. cDNA from fetal day 13.5 thymus was used as template for PCR with degenerate primers and the PCR products were cloned and sequenced. Of the first 30 sequences analyzed, 13 match to 11 separate known homeobox genes. Seven others match to five separate homeobox genes with a pattern that suggests that they are related to but distinct from known homeobox genes. These genes are thus potential candidates for involvement in thymus/thymocyte development. We plan to first determine the pattern of expression of these genes by PCR and Northern blot analysis. Genes that show patterns of specific interest (i.e., high or restricted expression in the thymus) will then be considered for further analysis by genetic approaches such as gene-targeting.

Publications:

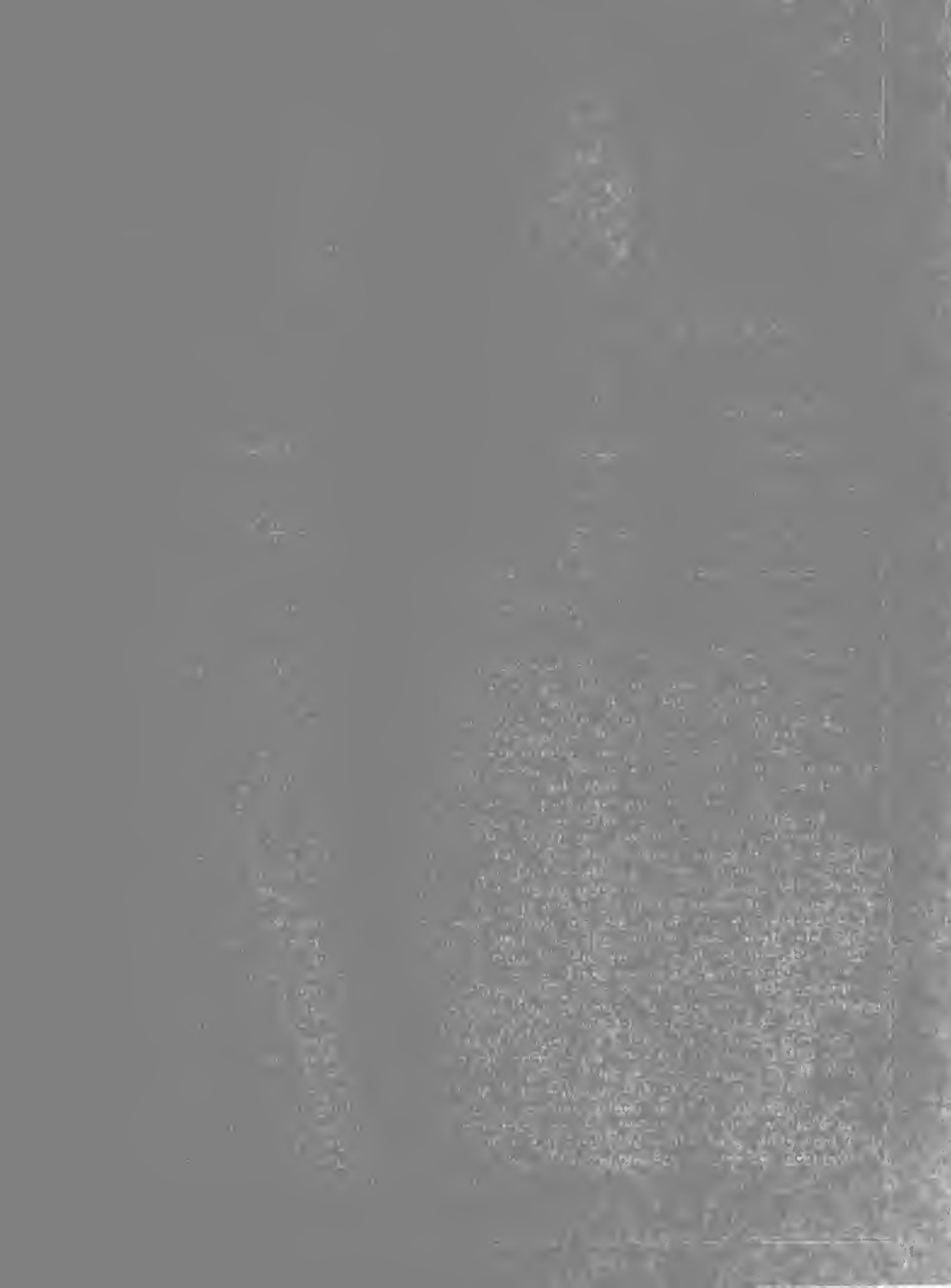
Cao X, Shores EW, Hu-Li J, Anver MR, Kelsall BL, Russell SM, Drago J, Noguchi M, Grinberg A, Bloom ET, Paul WE, Katz SI, Love PE, Leonard WJ. Defective lymphoid development in mice lacking expression of the common cytokine receptor γ chain. *Immunity* 1995;2:223-238.

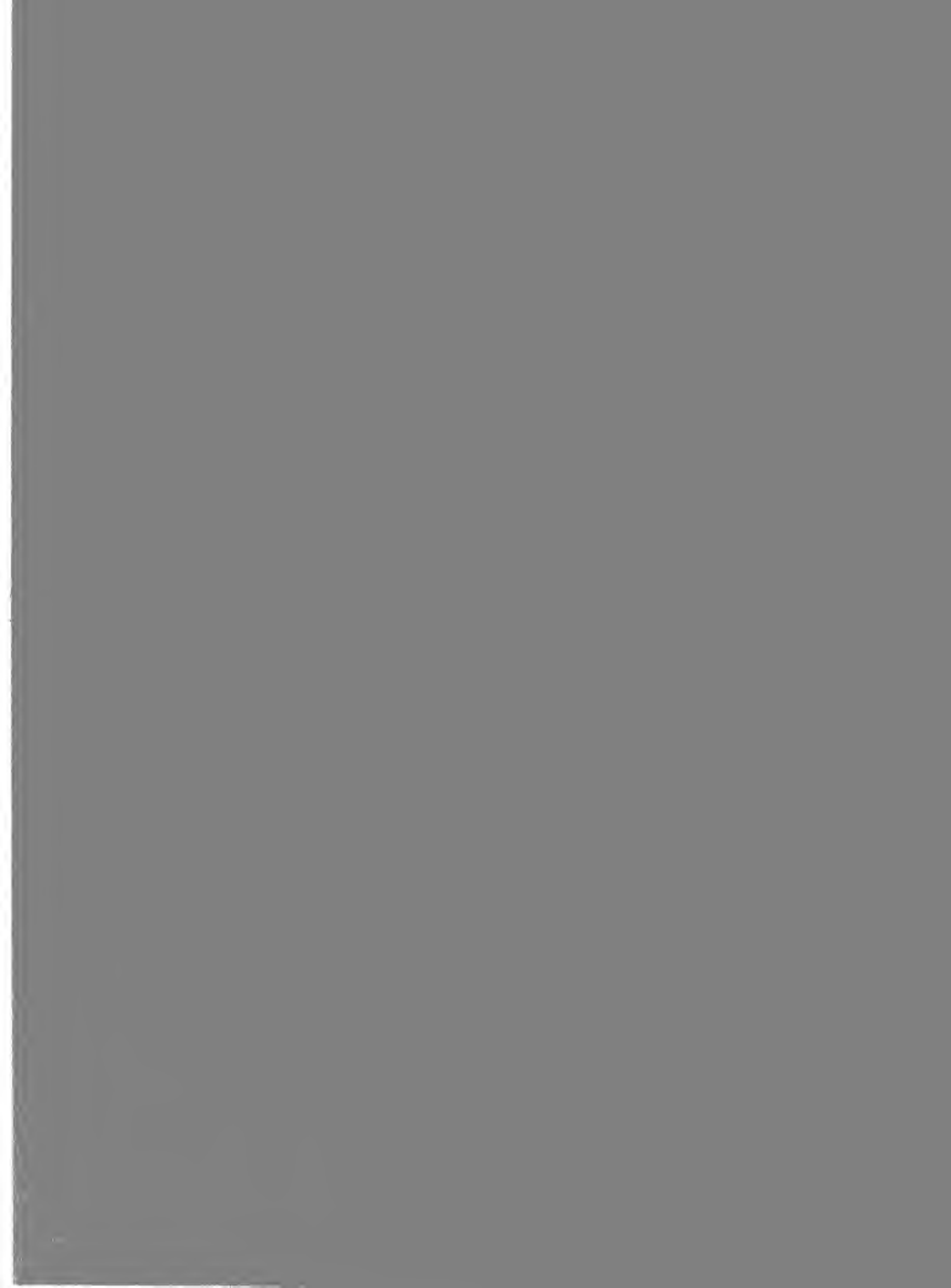
Leonard WJ, Shores EW, Love PE. Role of the common cytokine receptor γ chain in cytokine signaling and lymphoid development. *Immunol Rev*, in press.

Shores EW, Huang K, Tran T, Lee E, Grinberg A, Love PE. Role of TCR ζ chain in T cell development and selection. *Science* 1994;266:1047-1050.

Shores EW, Love PE. Insights into T cell development and signal transduction provided by TCR- ζ chain deficient mice. *Int Rev Immunol*, in press.

Sommers CL, Huang K, Shores EW, Grinberg A, Charlick DA, Kozak CA, Love PE. Murine *txk*: a protein tyrosine kinase gene regulated by T cell activation. *Oncogene* 1995;11:245-251.





DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER
 Z01 HD 00412-08 LMGR

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Molecular Regulation of Gene Expression

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute)

P.I.:	Richard J. Maraia	Senior Research Investigator	LMGR, NICHD
Other:	D.-Y. Chang	Intra Personnel Act	Univ.MD/LMGR, NICHD
	J. Sarrowa	IRTA	LMGR, NICHD
	A. Sakulich	Biologist	LMGR, NICHD
	J. Goodier	Visiting Fellow	LMGR, NICHD

COOPERATING UNITS (if any)

NCHGR/NIH (Eric Green)

LMGR-FY95

LAB/BRANCH

Laboratory of Molecular Growth Regulation

SECTION

Section on Human Cell Genetics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892-2753

TOTAL STAFF YEARS:

4.75

PROFESSIONAL:

3.75

OTHER:

1

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Research on small RNA gene expression was continued. RNA polymerase (pol) III is responsible for synthesizing transcripts such as tRNA and 5S rRNA which are required for protein synthesis, in addition to certain other small cytoplasmic transcripts such as Alu RNAs and Y RNAs. Investigation of mechanisms of pol III transcription and associated activities was continued. Alus are small retro-transposed elements endogenous to the human genome. About one million dispersed Alu elements constitute about 5% of the genetic material. Alu transposition, which depends on expression of Alu RNA, causes de novo genetic variability and insertional mutations in humans. Investigation of Alu RNA and Alu RNA binding proteins continued to yield insight into their complex regulation and propensity for retrotransposition. Y RNA-containing Ro RNPs are autoimmune antigens in humans and are targeted in autoimmune-mediated congenital complete heart block in infants. Summary of Major Findings: 1. SRP14-Alu RNA binding protein is associated with Alu transcripts in adenovirus-infected cells. This protein is overexpressed relative to other SRP subunits in primate cells. Overexpression of this protein is associated with the presence of a trinucleotide repeat (TNR) in the coding region of the SRP14 genes of higher primates. The SRP14 TNR mutation in the SRP14 gene coincided with the evolutionary events that produced a very large deposition of Alu retro-transposed elements into the genomes of higher primates. 2. The human autoimmune antigen known as La is a transcription termination factor. La controls reinitiation of small RNA genes by also functioning as an (re)initiation cofactor for RNA polymerase III. This activity augments La's activity as a termination factor to recycle both pol III and stable transcription complexes. Efficient recycling is necessary for production of the high levels of small RNAs that accumulate in vivo such as tRNA and 5S rRNA. 3. The human Y RNA gene family was cloned and its position mapped. This gene family represents a unique locus of four distinct pol III transcribed genes and resides in the vicinity of the telomeric end of human chromosome 7q36. The linear order of the single copy hY RNA genes was determined in terms of centromere-to-telomere orientation. Pol III-responsive upstream control elements of hY RNA genes were characterized.

Project DescriptionObjectives:

To understand the mechanisms that regulate small RNA expression in response to metabolic needs.

Methods Employed:

Tissue culture (TC) cells were used preparatively as a source of nuclear extract for *in vitro* transcription and RNA-binding protein. Isolation and analysis of active RNA polymerase III-containing transcription complexes utilized biotin-streptavidin-based solid state methods. *In vitro* mutagenesis followed by affinity purification of "tagged" recombinant proteins produced in bacteria was used to generate mutant protein for analysis in RNA binding, transcription complex reconstitution, and transcription termination assays. Yeast were transformed with mammalian class III genes and the small RNA products detected. Mammalian RNA-binding proteins were overexpressed in yeast. TC was used for gene mapping by Northern and Southern blot analyses of human X rodent somatic cell hybrids (SCH). cDNA was constructed from HeLa and owl monkey poly(A)+ RNA and isolated by PCR. Human gene isolation was accomplished by first screening yeast artificial chromosome (YAC) libraries. Gene mapping to specific human chromosomes was by analyses of: somatic cell hybrids, YAC contigs, and fluorescent *in situ* hybridization (FISH; outside service) of metaphase chromosomes. PCR-based methods were also used for site directed mutagenesis and generation of DNA templates (and mutants thereof) for T7 RNA polymerase-directed *in-vitro* RNA synthesis. RNA and ribonucleoprotein (RNP) analyses were also by primer extension and immunoprecipitation. RNA structural analyses was by native and denaturing gel electrophoresis, computer-assisted free energy, phylogenetic modeling, and nuclease probing. RNP analyses after reconstitution used electrophoretic mobility shift assay (EMSA), UV crosslinking and sucrose gradient analyses, and was assisted by immunoblot and immunoprecipitation. Dissociation constant determinations of interactions between pure RNA and pure proteins was by EMSA followed by computer-modeled (LIGAND) Scatchard analysis. Protein purification was by ion-exchange, affinity, and gel filtration chromatography; for amino acid sequence analysis purification was supplemented by SDS PAGE, proteolytic digestion, HPLC separation, and microsequencing.

Major Findings:**Alu RNA expression and the Alu RNA-binding protein.**

SRP14. Alu RNA binding protein as identified in the UMCB by an Alu RNA binding assay, purified to homogeneity, and cDNA-cloned, is now accepted to be identical to the signal recognition particle (SRP) polypeptide subunit known as SRP14 (1). Compounded modifications of the SRP14 gene occurred during early primate evolution. These included gene duplication, expansion of a trinucleotide repeat region (TNR) in the 3'-coding region of the active gene, and increased expression. The net result was the replacement of the ancestral SRP14 polypeptide with a more highly expressed variant of this Alu RNA binding protein. The human SRP14-homologous Alu RBP can bind to and presumably regulate Alu RNA metabolism (1). The multiple changes in the SRP14 gene occurred during the period which encompassed a high rate of Alu amplification which led to the insertion of the great majority of the Alu repeats inherited by the human genome. Because there is a coincidence in evolutionary time between the appearance of polymorphisms in this protein, its increased expression, and a major burst of Alu activity, it is reasonable to speculate that this protein may have been one determining factor in the great Alu proliferation that occurred during establishment of our ancestral primate genome.

The human gene represented by a single ~3.3 kbp *Pst* I fragment from chromosome 15q22 actively produces the ~18 kDa SRP14 protein (i.e. formerly known as Alu RBP) (1). We reported on the existence of a second SRP14-homologous sequence on human chromosome 11 (2). A human YAC library was screened for the chromosome 15 and chromosome 11 sequences and clones were obtained and verified. The chromosome 11 sequence appears to be a retroposed pseudosequence since its restriction map is similar to SRP14 cDNA (as opposed to genomic DNA) and because it can be PCR amplified as a single ~750 bp (cDNA-size) fragment with no evidence of introns. By contrast, the YAC clone corresponding to the chromosome 15 gene does not amplify by PCR (presumably because it is too large) and its restriction pattern is consistent with having introns as expected. Additional results reveal the existence of two other SRP14-homologous sequences in the human genome.

SRP9. The SRP-associated protein which binds to the Alu-homologous region of 7SL RNA is composed of two polypeptides; SRP9 and SRP14 (3). Although it was assumed and expected from indirect evidence that SRP9 was involved in the Alu RNA binding EMSA activity this has now been confirmed (4). SRP9 cDNA was cloned; it contains 0.25 kb of coding region and 1.3 kb of 3' UTR (4). Its 3' UTR was used as a highly specific probe to examine human DNA. Two genes appear to exist; preliminary data maps one definitely to chromosome 10 and the other probably to chromosome 1.

Function of SRP and possible function of Alu RNPs. SRP in one form or another appears to be a constituent of all cells. Mammalian SRP consists of three domains each of which confers a specific activity to the particle. Signal sequence recognition is mediated by the SRP54 subunit which then brings the Alu domain into close proximity to the ribosome thereby arresting further translation thereby preventing premature folding. This frozen complex (ribosome/mRNA/N-terminal polypeptide/SRP) is then transported to and docked at the endoplasmic reticulum and translation arrest is relieved. Translational elongation then serves to thread the linear nascent polypeptide through the E.R. pore. Thus, the three activities of SRP are signal sequence recognition, translation elongation arrest, and E.R. docking, and each is conferred by a distinct domain of a modular SRP. For example, removal of the Alu domain of SRP leaves the targeting activities intact (5, 6). SRP RNA (aka 7SL in mammals) serves as scaffold onto which its six associated polypeptides are assembled in specific orientation. This view of SRP RNA attributes SRP activities entirely to the polypeptides although the possibility that SRP RNA is more "active" than simply providing a scaffold remains a possibility.

Experiments to investigate the possibility of an active role for SRP RNA in SRP activity have begun. Certain highly conserved residues in SRP RNA which overlap with a SRP9/14 binding site are being examined for their role in SRP structure and function. Preliminary results suggest a dynamic structure in the Alu domain of SRP RNA which is differentially recognized by SRP9/14. The role of the nucleotide residues involved in this feature of SRP will be examined with regard to translation elongation arrest.

The Alu domain-associated SRP9 kDa and SRP14 kDa polypeptides form a stable heterodimer in the absence of RNA. In SRP this confers translational control on a subset of ribosome-associated mRNAs yet a role for these proteins acting independently of SRP has not been explored (3, 5, 6, 7). Our results indicate that (i) SRP9/14 is present at 10-20 fold molar excess relative to other SRP subunits including SRP (7SL) RNA (2). The presence of Alu sequences in hnRNA as well as many mRNAs suggests the possibilities that this protein may affect Alu-containing RNA metabolism and/or translatability (1). Previous data demonstrated a ~5 fold increase in steady state levels of scAlu RNA in response to a ~5 fold increase in Alu RBP (1). Alu transcripts are induced 20-50 fold in Adenovirus infected cells (8, 9, 10). Potential roles for Alu RNPs in viral-induced reprogramming of the translational apparatus are being considered. Experiments to characterize cell biological aspects of Alu RNA induction are underway. Recent data reveals that scAlu and full

length Alu RNA are associated with SRP9/14 as small RNPs in infected as well as uninfected cells, and that these are induced by adenovirus infection. (manuscript in preparation).

La. In addition, full length cytoplasmic Alu transcripts are also associated with La as well (as expected for all nascent transcripts of pol III). La has been implicated in viral gene expression at the translational level (11, 12, 13) including HIV (14, 15). Work from several laboratories indicate that La is an abundant protein ($\sim 2 \times 10^7$ copies/cell) that is apparently targeted by a variety of viral mRNAs during infection. In collaboration with Dr. K.-T. Jeang (NIAID) it was found that HIV infection of tissue culture cells did not induce scAlu or full length Alu RNA; unpublished. In addition, owing to its role as a transcription termination/transcript maturation factor (16), La may regulate the levels of key tRNAs in the cell and may indirectly affect translation.

One group reported that La may have the ability to regulate the double strand RNA activated RNA-binding protein, PKR in vitro (17). However, in collaboration with T. Dever (LMG) it was found that overexpression of human La had no effect on human PKR activity expressed in yeast cells (unpublished).

Transcription by RNA polymerase III: La

Much has been learned from the study of pol III-transcribed (class III) genes. It has recently been established that different class III genes use different promoter structures to assemble initiation complexes by different pathways (reviewed in (18)). Studies of transcription factor utilization by class III genes has been a long time model for other eukaryotic RNA polymerases, as the first eukaryotic genes transcribed in vitro were pol III templates (19).

Once formed, pol III transcription initiation complexes remain stable through multiple rounds of transcription. In addition to the possibility that the stability of these complexes might be a means for the genome to determine and maintain the differentiated state of certain class III genes (20), transcription complex stability makes sense in terms of efficient small RNA production such as is required for protein synthesis (tRNA) and ribosome assembly (5S rRNA). Pol III transcription is regulated in a cell cycle-dependent manner and is probably coordinated with RNA polymerases I and II during cell proliferation (21, 22, 23). It is probably an important feature of class III genes that they do not have to disassemble and reassemble transcription complexes for each molecule of RNA synthesized, especially since most class III genes are expressed at very high levels. For example, the 7SL RNA component of SRP is present at about ten million copies per cell but only about ten copies of the 7SL RNA gene reside in diploid cells. Thus, each 7SL gene must generate about one million RNA products per cell doubling. Calculations based on HeLa cell proliferation yield rates of ~ 100 transcripts per gene per minute. A similar rate can be calculated for 5S rRNA. A need for high level reinitiation efficiency may be a specialized function of class III complexes because unlike protein-encoding genes whose products can be amplified by translation and then again by catalytic activity of an enzyme, most pol III products are structural and are used directly. In any case, de novo assembly of initiation complexes on class III genes determines efficiency of the first round of transcription whereas all subsequent cycles would be due to accurate reinitiation, elongation and termination. Some class III genes are so short (eg. tRNA) that they can not accommodate more than one or two elongating polymerases per RNA length; this is in sharp contrast to pol I and pol II templates which may be thousands to hundreds-of-thousands of basepairs long and which can load many polymerases onto the transcription unit long before termination of the first. In addition, although pol I and pol II termination is tightly coupled with nascent transcript 3'-end processing (probably as a means to destabilize the transcription complex allowing dissociation of the polymerase), pol III termination appears to be more directly linked to polymerase release. Without accurate and efficient termination and transcript release, efficiency of class III RNA synthesis declines; for example in vitro studies suggest that stalling of pol III at the terminator without transcript release blocks subsequent rounds of RNA synthesis (16). Based on this reasoning and the necessity for efficient class III gene utilization, it might be expected that

safeguards to ensure efficient termination may be part of the pol III machinery. Recently, pol III terminators were shown to affect transcription efficiency in vivo (24).

Termination by pol III must also be accurate since the 3' ends of small RNAs are often important determinants of RNA structure. How an elongating pol III is converted to one which terminates efficiently and accurately to leave the template readily reinitiatable are unresolved issues. Understanding these aspects of RNA synthesis will require knowledge of the basic mechanisms involved in post-initiation activities of RNA polymerase III.

Class III genes use a simple, sequence specific signal in DNA to terminate RNA synthesis by pol III. This terminator corresponds to a stretch of dT residues in the strand of DNA that represents the transcript sequence (25), although the actual signal is probably in the dA tract in the template strand (26). The dT_{≥4} terminator is common to class III genes regardless of their promoter structure. In higher eukaryotes the pol III terminator consists of 4 or more dT residues. This termination signal is reminiscent of rho-independent terminators used by *E. Coli* RNA polymerase which contain about 8 dA residues in the template strand. Yeast pol III appears to require at least 6 or 7 dT residues for efficient termination (27).

Termination by pol III occurs in two steps; polymerase pausing at the dT_{≥4} signal, followed by disengagement of the polymerase and transcript release (28). The dT_{≥4} signal is recognized by purified pol III itself and is sufficient to cause the polymerase to terminate transcription, although some aspects of termination may be affected by pol III-associated factors in the transcription complex (16, 28).

The human autoimmune antigen/phosphoprotein known as La was initially identified as a small nuclear RNA-binding protein and later shown to participate in transcription termination by RNA polymerase (pol) III (16, 29, 30). La protein mediates transcription termination and transcript release from paused pol III transcription complexes. La is a ~50 kDa RNA binding protein that can associate with many different transcripts synthesized by pol III by binding to the oligo(rU) tract which corresponds to the pol III termination signal that defines the 3'-terminus of all pol III-synthesized nascent RNAs.

La's role in transcription was previously thought to be limited to a singular molecular event at termination. We recently used VA1 and other genes to examine reinitiation from transcription complexes that were systematically depleted of pol III initiation activity. This was possible due to the ability to reconstitute and then partially dismantle transcription complexes that have been assembled on solid state supports in test tubes. It was found that these pol III-depleted initiation complexes are incompetent for transcription by supplemental purified pol III unless La (highly purified as recombinant protein from bacteria) is also provided. La is necessary for a single reinitiation as well as multiple rounds of RNA synthesis from these initiation complexes. Thus, this novel (re)initiation factor activity of La appears to be distinct from and augments its activity as a termination factor. La is necessary for efficient use of pol III and preassembled transcription complexes and is an activator of class III gene transcription.

The RNA-binding domain of La resides within its first 235 amino acids, while the remaining ~200 aa domain has no function attributed to it and is therefore implicated in protein-protein interactions. We are presently examining the structure/function aspects of the La protein in terms of its ability to mediate termination as well as reinitiation by pol III. Thus, the ultimate proof of distinct and separable activities of La at termination and reinitiation must await future analyses of deletion and substitution mutants of the La protein in the appropriate assays carried out in the UMCB.

Human Y scRNA gene cloning, upstream regulation, and chromosomal mapping. Four different small cytoplasmic (sc)RNAs designated hY1, hY3, hY4 and hY5 accumulate in human cells as ribonucleoproteins (RNP) associated with the Ro and La autoantigen proteins (31,

32). *In vivo* Y RNAs are most abundant in heart and brain tissues (see ref. (33), and this pattern is established during fetal development (UMCB, unpublished data). The presence of autoimmune antibodies directed against Ro RNPs is tightly associated with the development of congenital complete heart block in infants born to anti-Ro positive mothers whose autoantibodies passively cross the placenta (34). Furthermore, it has been reported that these anti-Ro autoantibodies immunohistolocalize to the affected infants' cardiac conduction system, and in experimental systems inhibit cardiac cell membrane repolarization (35, 36) and refs. therein). Although initially identified as human autoantigens Ro Y RNPs have since been found in several species including *Xenopus* and *c. elegans* (37, 38, 39, 40, 41). Unlike most pol III synthesized RNAs, a subset of mammalian Y RNAs retain their oligo(U) 3' termini and remain associated with the La antigen, although the significance of this if any is unknown. Nonetheless, in this regard Y RNAs are similar to the adenovirus encoded VA_I and VA_{II} RNAs and the Epstein Barr virus small RNAs which are synthesized by pol III and remain complexed with La; each of these virally-encoded class III templates produces RNAs which inhibit the double strand-activated kinase that regulates eIF2 α . The evolutionary conservation of Ro Y RNPs coupled with their regulated expression and involvement in human pathologic conditions which are manifest in development provide ample evidence to indicate an important biological role for these RNPs even though no function has yet been ascribed to them (31, 32, 34, 37, 38, 39, 40, 42). Little is known about their expression even though Y RNAs unlike most other class III genes appear to represent tissue-selective products.

The first Y RNA genes cloned were hY1 and hY3 both of which reside within 4 kb of each other (32). However, further attempts by several groups to clone mammalian Y RNA genes have been impeded by the presence of an abundance of dispersed Y pseudogenes in mammalian genomes (39, 43, 44). The UMCB reported the cloning of hY4 RNA gene (45). Recently, and again in collaboration with Eric Green (NCHGR), the remaining human Y RNA gene, hY5 was cloned. Like hY4, the hY5 scRNA gene is also dependent on its upstream control elements for transcription in transfected cells (45, 46).

The linear order of the hY RNA genes on chromosome 7q36 was determined to be: centromere-hY5-hY4-hY3/hY1-telomere (hY3 and hY1 genes are within 4 kb of each other and their orientation relative to each other has not yet been determined). This gene family appears to represent a unique locus of upstream-regulated class III genes that are differentially regulated. Investigation of the regulation of this gene family is expected to reveal novel mechanisms of pol III regulation.

Significance to biomedical research and the program of the Institute.

The impact of Alu SINE amplification on the human genome has been substantial. Germ line Alu mobility has caused recognizable hereditary disorders in humans by mechanisms that are modeled but poorly understood. Alu transcripts mediate Alu retrotransposition yet factors involved have not previously been identified. The primate SRP14-Alu RBP is significantly larger than its mammalian counterparts and is overexpressed specifically in primates. SRP14-Alu RBP modulates Alu RNA levels *in vivo* and appears to have been involved in the primate-specific amplification of Alu elements which shaped the human genome. Thus, this protein is an excellent candidate factor involved in mediating Alu transposition; however an experimental system for studying this is still not yet available. In addition, many human mRNAs contain Alu sequences in their untranslated regions; an unpredictably high percentage of these are related to immune and or cell recognition suggesting that Alu elements have contributed to their functional genetic variability. The facts that SRP14 is involved in modulating translation of secreted proteins and that Alu RNPs increase dramatically in adenovirus-infected cells suggests an independent role for Alu in response to viral infection.

Pol III must produce very high levels of tRNAs and 5S rRNAs required for protein synthesis. Understanding the factors, mechanisms and regulation of the efficient use of limited amounts of pol III in cells is our goal. Elucidation of regulated physiologic connections between pol III activity, protein synthesis and DNA replication should benefit our understanding of basic aspects of cellular function.

Y RNA gene function in humans (or any species) is completely unknown yet hY RNA-containing Ro RNPs are targeted by autoimmune disorders of adults and infants. Cloning and functional characterization of the hY4 and hY5 RNA genes which each produce pol III transcript constituents of scRNPs was accomplished.

These advances were directed toward achieving the goals of the program of the institute.

References

1. D. Y. Chang, *et al.*, *Mol. Cell Biol.* **14**, 3949 (1994).
2. D. Y. Chang, N. Sasaki-Tozawa, L. K. Green, R. J. Maraia, *Mol. Cell Biol.* **15**, 2109 (1995).
3. K. Strub and P. Walter, *Mol. Cell Biol.* **10**, 777 (1990).
4. K. Hsu, D. Y. Chang, R. J. Maraia, *J. Biol. Chem.* **270**, 10179 (1995).
5. V. Siegel and P. Walter, *Nature* **320**, 81 (1986).
6. V. Siegel and P. Walter, *Cell* **52**, 39 (1988).
7. K. Strub, J. B. Moss, P. Walter, *Mol. Cell Biol.* **11**, 3949 (1991).
8. B. Panning and J. R. Smiley, *Mol. Cell Biol.* **13**, 3231 (1993).
9. B. Panning and J. R. Smiley, *Virology* **202**, 408 (1994).
10. V. R. Russanova, C. T. Driscoll, B. H. Howard, *Mol. Cell Biol.* **In Press**, (1995).
11. K. Meerovitch, J. Pelletier, N. Sonenberg, *Genes Dev.* **67**, 3798 (1989).
12. Y. V. Svitkin, *et al.*, *J. Virol.* **68**, 1544 (1994).
13. K. Meerovitch, *et al.*, *J. Virol.* **67**, 3798 (1993).
14. Y.-N. Chang, D. J. Kenan, J. D. Keene, A. Gaignol, K.-T. Jeang, *J. Virol.* **68**, 7008 (1994).
15. Y. V. Svitkin, A. Pause, N. Sonenberg, *J. Virol.* **68**, 7001 (1994).
16. R. J. Maraia, D. J. Kenan, J. D. Keene, *Mol. Cell Biol.* **14**, 2147 (1994).
17. Q. Xiao, *et al.*, *Nucl. Acids Res.* **22**, 2512 (1994).
18. N. Hernandez, in *Transcriptional Regulation* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1992), pp. 281.

19. G.-J. Wu, *Proc. Natl. Acad. Sci. USA* **75**, 2175 (1978).
20. D. F. Bogenhagen, W. M. Wormington, D. D. Brown, *Cell* **28**, 413 (1982).
21. P. Hartl, J. Gottesfeld, D. J. Forbes, *J. Cell Biol.* **120**, 613 (1993).
22. J. M. Gottesfeld, V. J. Wolf, T. Dang, D. J. Forbes, P. Hartl, *Science* **263**, 81 (1994).
23. R. J. White, T. M. Gottlieb, C. S. Downes, S. P. Jackson, *Mol. Cell. Biol.* **15**, 1983 (1995).
24. W. M. Chu, W. M. Liu, C. W. Schmid, *Nucleic Acids Research* **23**, 1750 (1995).
25. D. F. Bogenhagen and D. D. Brown, *Cell* **24**, 261 (1981).
26. X. Wang and W. R. Folk, *J Biol Chem* **269**, 4993 (1994).
27. R. A. Koski, S. G. Clarkson, J. Kurjan, B. D. Hall, M. Smith, *Cell* **22**, 415 (1980).
28. F. E. Campbell and D. R. Setzer, *Mol. Cell. Biol.* **12**, 2260 (1992).
29. E. Gottlieb and J. A. Steitz, *EMBO J.* **8**, 851 (1989).
30. E. Gottlieb and J. A. Steitz, *EMBO J.* **8**, 841 (1989).
31. J. P. Hendrick, S. L. Wolin, J. Rinke, M. R. Lerner, J. A. Steitz, *Mol. Cell. Biol.* **1**, 1138 (1981).
32. S. L. Wolin and J. A. Steitz, *Cell* **32**, 735 (1983).
33. S. L. Wolin and J. A. Steitz, *Proc. Natl. Acad. Sci.* **81**, 1996 (1984).
34. P. V. Taylor, *et al.*, *New England Journal of Medicine* **315**, 667 (1986).
35. J. S. Deng, L. W. J. Bair, S. Shen-Schwarz, R. Ramsey-Goldman, T. J. Medsger, *Arthritis Rheum* **30**, 1232 (1987).
36. E. Alexander, J. P. Buyon, T. T. Provost, T. Guarnieri, *Arthritis & Rheumatism* **35**, 176 (1992).
37. Y. Itoh, J. D. Kriet, M. Reichlin, *Arthritis & Rheumatism* **33**, 1815 (1990).
38. M. R. Lerner, J. A. Boyle, J. A. Hardin, J. A. Steitz, *Science* **211**, 400 (1981).
39. G. J. Pruijn, P. A. Wingens, S. L. Peters, J. P. Thijssen, W. J. van Venrooij, *Biochim Biophys Acta* **1216**: 395-401 **1216**, 395 (1993).
40. C. A. O'Brien, K. Margelot, S. L. Wolin, *Proc. Natl. Acad. Sci. USA* **90**, 7250 (1993).
41. D. J. Van Horn, D. Eisenberg, C. A. O'Brien, S. L. Wolin, *RNA* **In press**, (1995).
42. C. A. O'Brien and J. B. Harley, *EMBO J.* **9**, 3683 (1990).

43. D. Crouch and E. H. H. Liebke, *Nucl. Acids Res.* **17**, 4890 (1989).
44. C. A. O'Brien and J. B. Harley, *Gene* **116**, 285 (1992).
45. R. J. Maraia, N. Sasaki-Tozawa, C. T. Driscoll, E. G. Green, G. J. Darlington, *Nucl Acids Res* **22**, 3045 (1994).
46. R. J. Maraia, E. Brinkman, A. L. Sakulich, E. D. Green, *In preparation* (1995).

Proposed Course

1. Continue investigation of regulation of class III gene expression by the autoantigen/phosphoprotein La. Initiate study of La-mediated responses to intracellular metabolic stimuli.
2. Examine role of SRP and SRP-homologous Alu RBP in adenoviral infection.
3. Initiate investigation of structure/function relationship of Alu domain of SRP.
4. Continue to characterize human SRP14-homologous Alu RBP gene on chromosome 15q22, and related sequence on chromosome 11. Continue characterization of TNR within Alu RBP gene in individuals of different ethnic/geographical extraction.
5. Refine the physical map of Y RNA gene family on chromosome 7q36 including more refined relationship to telomere.

Publications

Chang DY, Sasaki-Tozawa N, Green LK, Maraia RJ. A Trinucleotide repeat-associated increase in the level of *Alu* RNA-binding protein occurred during the same period as the major *Alu* amplification that accompanied anthropoid evolution. *Mol Cell Biol* 1995;**15**:2109-16.

Hsu K, Chang DY, Maraia RJ. Human signal recognition particle (SRP) Alu-associated protein also binds Alu interspersed repeat sequence RNAs: characterization of human SRP9. *J Biol Chem* 1995;**270**:10179-86.

Maraia RJ, Sasaki-Tozawa N, Driscoll CT, Green EG, Darlington GJ. The human Y4 small cytoplasmic RNA gene is controlled by upstream elements and resides on chromosome 7 with all other hY scRNA genes. *Nucl Acids Res* 1994;**22**:3045-52.

Maraia RJ. ed. Impact of short interspersed elements (SINEs) on the host genome, Austin: R. G. Landes, 1995.

Maraia RJ, Sarrowa J. Alu-family SINE RNA: Interacting proteins and pathways of expression. In: Maraia RJ. ed. Impact of short interspersed elements (SINEs) on the host genome. Austin: R. G. Landes, 1995;163-196.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 00505-02 LMGR

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Eukaryotic Transcriptional Regulation

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute)

P.I.:	Yoshihiro Nakatani	Visiting Associate	LMGR, NICHD
Other:	Tetsuro Kokubo	Visiting Fellow	LMGR, NICHD
	Ho, Chi-Yip	Visiting Fellow	LMGR, NICHD
	Junichi Nishikawa	Special Volunteer	LMGR, NICHD
	Yang, Xiang-Jiao	Visiting Associate	LMGR, NICHD
	Ogryzko, Vasily V.	Visiting Associate	LMGR, NICHD

COOPERATING UNITS (if any)

LAB/BRANCH

Laboratory of Molecular Growth Regulation

SECTION

Section on Human Cell Genetics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

4.25

PROFESSIONAL:

4.25

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The group is focused on the mechanisms of 1) how gene-specific transcriptional activators interact with RNA polymerase II transcription machinery to regulate messenger RNA synthesis; and 2) how tumor viral proteins perturb normal cellular interactions for transforming the cells.

Transcription initiation factor TFIID plays a central role in transcriptional regulation by facilitating promoter responses to various activators. TFIID is a large multisubunit complex, consisting of the TATA-box binding protein (TBP) and a number of tightly associated subunits (TAFs). We observed that both *Drosophila* TAFII230 (dTAFII230) and the homologous yeast TAFII145 (yTAFII145) interact directly with TBP and inhibits TBP binding to the TATA box in the absence of other TAFs. Deletion analysis indicates that N-terminal regions of TAFs (156 residues in dTAFII230; 71 residues in yTAFII145) have both binding and inhibitory activities. These regions contain two functional subdomains, I and II, which synergistically contribute to stable TBP-binding. Importantly, subdomain-II recognizes basic repeats on the upper surface of TBP that are also critical for TFIIA-binding, whereas subdomain-I competes with transactivator VP16 on the under surface of TBP. To determine the biological properties of the inhibitory activity, the yTAFII145 gene was replaced with the mutant lacking the inhibitory domain. The mutant strains are viable, but grow slowly. These results imply that the inhibitory domain plays a key role in transcriptional regulation by communicating with TFIIA and activators.

The cellular protein p300/CBP is a target of the adenoviral E1A oncoprotein. Although this interaction is crucial for E1A transforming function, the mechanisms by which E1A modulates cell growth through p300/CBP are poorly understood. We describe here a cellular CBP-associated factor (CAF) which binds to the CBP site recognized by E1A. Although CAF was cloned as a human counterpart of the yeast transcriptional cofactor GCN5, the nonconserved N-terminal region of CAF mediates interaction with CBP in vitro. This interaction is disrupted by E1A. Remarkably, overexpression of CAF in HeLa cells inhibits cell cycle progression. Consistent with the in vitro competition, this effect is partially counteracted by E1A. These results provide new insights into the molecular mechanisms of tumorigenesis.

Project DescriptionObjectives:

To understand the mechanisms of 1) how gene-specific transcriptional activators interact with RNA polymerase II transcription machinery to regulate messenger RNA synthesis; and 2) how tumor viral proteins perturb normal cellular interactions for transforming the cells.

Methods Employed:

General methods in molecular biology and protein chemistry were employed.

Major Findings:1. Mechanism of Transcriptional activation**Introduction**

Transcription of protein-encoding genes in eukaryotes is regulated by various gene-specific transcriptional factors that bind to distinct DNA control elements. Although these factors are thought to regulate transcription through interactions with the RNA polymerase II transcription machinery, their mechanism of action is poorly understood. In addition to RNA polymerase II, accurate transcription initiation in vitro with natural components requires at least six general initiation factors: TFIIA, TFIIB, TFIID, TFIIIE, TFIIF, and TFIIH. The first step in preinitiation complex formation involves TFIID binding to the TATA box in the core promoter, an interaction which may be facilitated by TFIIA. The resulting complex is then recognized by TFIIB to form a complex which in turn recruits the preformed RNA polymerase II/TFIIF complex to the promoter. The subsequent binding of TFIIIE and TFIIH completes formation of the preinitiation complex. These assembly steps are potential targets for regulatory factors, which through interactions with the corresponding general factors may alter the rate of preinitiation complex formation or function.

TFIID is a multimeric protein complex consisting of the TATA box binding subunit (TBP) and many of tightly associated polypeptides (TAFs). While the native TFIID complex mediates activator- (or repressor-) regulated transcription in conjunction with other components of transcription machinery, the derived TBP by itself mediates only basal transcription. Thus, it is believed that one or more TAFs are essential for transmitting signals from various activators (or repressors) to the remainder of the general transcriptional machinery.

We have identified at least 9 polypeptides (p230, p110, p85, p62, p42, p28a, p28b, p22, and TBP) as *Drosophila* TFIID subunits on the basis of an analysis of derived cDNA products. The largest subunit, dTAFII230, shows significant sequence similarity with the product of human cell cycle gene (CCG1), also identified as the largest subunit of human TFIID complex. The fact that CCG1 mutations arrest cells in late G1 and inhibit transcription only of specific genes suggests a specific role for a part of CCG1 in cell cycle regulated transcription. A more direct role for dTAFII230 in the regulation of initiation is indicated by the ability of dTAFII230 to inhibit both TATA box binding and basal transcription activities of TBP in the absence of other TFIID subunits. Determining how dTAFII230 negatively regulates TBP function represents an important step toward understanding the various mechanisms which regulate transcription initiation. For this purpose, yeast genetics could be a powerful way to determine both in vitro and in vivo function of the negative interaction. Thus we have cloned the homologous yeast TAFII145 (yTAFII145) and analyzed in detail.

TAF Competes with Positive Regulators on TBP Surfaces.

Both dTAFII230 and yTAFII145 recombinant polypeptides form stable complexes with TBP and inhibits TBP function. The N-terminal portions (156 residues in dTAFII230; 71 residues in yTAFII145) have both binding and inhibitory activities although this region is conserved extremely poorly between the species. A detailed analysis indicates that this portion contains two functional subdomains, I and II, which synergistically contribute to stable TBP-binding. Importantly, subdomain-II recognizes basic repeats on the upper surface of TBP that are also critical for TFIIA-binding, whereas subdomain-I competes with transactivator VP16 on the under surface of TBP. These results imply that their inhibition of TATA box-binding could be modulated by positive factors VP16 and TFIIA. To determine the biological properties of the inhibitory activity, the yTAFII145 gene was replaced with the mutant lacking the inhibitory domain. The mutant strains are viable, but growth slowly. These results suggest that the inhibitory domain mainly contributes to regulatory functions, and may not be required for the formation of multimeric TFIID complex.

2. Mechanism of Transformation by Tumor Viral Factors.

Introduction

The transforming proteins encoded by adenovirus and several other small DNA tumor viruses disturb host cell growth control by interacting with cellular factors that normally function to repress cell growth. One of the most intensively studied of these viral proteins, the product of the adenovirus E1A gene, is itself sufficient for transformation. E1A transforming activity resides in two distinct domains, the targets of which are the cellular factors p300 and the retinoblastoma susceptibility gene product (RB). Interactions of E1A with p300 and RB are thought to influence functionally distinct growth regulatory pathways, allowing the two domains to contribute additively to transformation.

Recently, p300 was found to be very closely related to CBP, a factor that binds selectively to the protein kinase A-phosphorylated form of CREB. CBP and p300 exhibit strong amino acid sequence similarity, and furthermore share the capacity to bind both CREB and E1A. Although p300/CBP by itself cannot bind to DNA, it is recruited to promoter regions via interaction with sequence-specific activators including CREB, Jun and YY1, and is thus able to activate transcription. E1A negatively regulates these activators via its interaction with p300/CBP. Importantly, E1A does not appear to disrupt p300/CBP-activator interactions; hence it is crucial to determine a cellular counterpart of E1A which binds to p300/CBP.

Cloning of CBP Associated Factor (CAF).

We sought to isolate cofactors for p300/CBP based on analogy between species. CBP binds to Jun in a phosphorylation-dependent manner in association with stimulation of transcription. In yeast, GCN4 is believed as a Jun counterpart on the basis of similarities in DNA recognition as well as the participation of both proteins in UV signaling pathways. Yeast genetic screening has allowed isolation of various cofactors for GCN4, including GCN5, ADA2 and ADA3. These factors are considered to function as a complex (or, alternatively, in a closely related pathway) based on genetic and protein-protein interaction studies. From these observations we hypothesized that a human counterpart of these factors may be involved in the Jun-p300/CBP pathway.

Comparison of GCN5 protein sequence with databases revealed significant similarity ($P=2.1 \times 10^{-48}$) with a randomly sequenced human cDNA. With use of this cDNA as a probe, we isolated cDNA encoding the full-length 93 kD protein. The deduced protein sequence displays a strong similarity (50% identity, 68% similarity) with GCN5 at the C-terminal region (residue 488-832) (Fig. 3). On the other hand, no obvious similarity with the N-terminal region was revealed by database searching. Interaction study revealed that the N-terminal region interacts with CBP.

Thus we termed the cDNA encoded polypeptide CBP-associated factor (CAF) rather than human GCN5.

CAF and E1A Competitively Bind to the Same CBP Regions

Interaction between CAF and various CBP fragments was tested. Two CBP segments, corresponding to amino acids 1,678-1,880 and 1,801-2,000, interacted specifically with CAF. Importantly, E1A also bound to these fragments. These results raises the possibility of direct competition between CAF and E1A for binding to CBP. To test this prospect, a competition experiment was performed with the use of partially purified proteins. As expected, CAF-CBP complex was disrupted by E1A. Thus we conclude that both CAF and E1A cannot bind to CBP simultaneously.

CAF represses cell cycle progression

The proposal that p300 is involved in the maintenance of the G0/G1 state is based on indirect evidence, namely, that the p300-binding domain of E1A is important for its transforming activity. This predicts that overexpression of p300 should have a negative effect on G1/S progression. We surmised that CAF, by binding to and forming a functional complex with p300, might also retard entry into S phase.

To address the above possibilities, we examined whether transient expression of either p300 or CAF would affect the rate of G1/S transit in HeLa cells. Both p300 and CAF negatively affected the distribution of cells between G1 and S phases in the above assay. Consistent with the in vitro competition, this effect is partially counteracted by E1A.

Proposed Course:

TAF project:

- 1) Test the genetic interactions between yTAFII145 and VP16 or TFIIA.
- 2) Isolation of suppressor mutants of yeast lacking the inhibitory domain.
- 3) How does the inhibitory domain contribute to transcriptional activation in yeast?

CAF project:

- 1) How does CAF regulates CBP-dependent transcription.
- 2) How does E1A inhibits CBP-dependent transcription.

To address these questions we will employ both transfection and in vitro transcription systems.

Publications

Gong DW, Mortin MA, Horikoshi M, Nakatani, Y. Molecular cloning of cDNA encoding the small subunit of Drosophila transcription initiation factor TFIIF, Nucl Acids Res 1993; 23:1882-6.

Hisatake K, Ohta T, Takada R, Guermah M, Horikoshi M, Nakatani Y, Roeder RG. Evolutionary conservation of human TATA-binding-polypeptide-associated factors TAFII31 and TAFII80 and interactions of TAFII80 with other TAFs and with general factors, Proc Natl Acad Sci USA 1995; in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE NOTICE OF INTRAMURAL RESEARCH PROJECT		PROJECT NUMBER Z01 HD 01310-09 LMGR
PERIOD COVERED October 1, 1994 to September 30, 1995		
TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.) Developmental Gene Regulation of the Immune System		
PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and Institute) PI: K. Ozato Head, SMGI LMGR, NICHD See Attached		
COOPERATING UNITS (if any) Laboratory of Immunopathology, NIAID (Dr. Herbert Morse).		
LAB/BRANCH Laboratory of Molecular Growth Regulation		
SECTION Section on Molecular Genetics of Immunity		
INSTITUTE AND LOCATION NICHD, NIH, Bethesda, Maryland 20892		
TOTAL STAFF YEARS: 11.2	PROFESSIONAL: 9.5	OTHER: 1.7
CHECK APPROPRIATE BOX(ES) <input type="checkbox"/> (a) Human subjects <input type="checkbox"/> (b) Human tissues <input checked="" type="checkbox"/> (c) Neither <input type="checkbox"/> (a1) Minors <input type="checkbox"/> (a2) Interviews		
SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.) <p> This group works on gene regulation in the immune system. Our focus has been to elucidate the role of two transcription factors, RXRβ and ICSBP that we have isolated previously. RXRβ is a member of the nuclear hormone receptor superfamily and by heterodimerizing with retinoic acid receptor (RAR), it regulates retinoic acid (RA) responsive genes. ICSBP is a member of the interferon regulatory factor (IRF) family and is involved in viral/interferon (IFN) mediated gene regulation. To further study the role of ICSBP and other members of the IRF family in the immune system, a dominant negative ICSBP has been constructed which has only the DNA binding domain (DBD) and lacks a large C-terminal domain. This construct was stably transfected into a human monocytic cell line U937 cells, and a number of clones expressing the DBD have been established. The DBD clones exhibited altered infectivity by several viruses: they were refractory to infection by vaccinia virus, ectromelia virus as well as human immunodeficiency virus-1. In contrast, control clones and those expressing the intact ICSBP were strongly infected by these viruses. The DBD clones were, however, infected by vesicular stomatitis virus (VSV) as efficiently as control clones. But, IFN pretreatment known to confer protection against VSV, completely failed to protect DBD clones. As expected control clones and those expressing intact ICSBP were fully protected from VSV infection by IFNs. These results indicate that some viruses depends on the IRF proteins for their infection and growth in host cells. By genomic footprinting analysis we have studied binding of RXR heterodimers to the promoter of an RA responsive gene (RARβ) in P19 EC cells. We show that the RA responsive element (RARE) in the RARβ gene is not occupied prior to RA treatment, but becomes occupied within 1 h of RA treatment, without requiring new protein synthesis. This in vivo occupancy accompanied factor binding to other elements in the promoter. The occupancy at all sites was rapidly reversed when RA was withdrawn from culture medium. On the other hand, P19 cells expressing a dominant negative RXRβ that blocks binding of endogenous RAR/RXR heterodimers to the RARE, did not exhibit occupancy at all elements in the promoter. Our observations underscore the significance of ligand for receptor/DNA interaction in vivo and demonstrate the presence of a hierarchy in factor binding processes. </p>		

Principle investigator
Keiko Ozato, Ph.D

Other Personnel:

Professional		Starting Year
Nisan Bhattacharyya	Ph.D NRC associate	1994
Jorge Blanco	Ph.D Visiting Fellow	1992
Christina Contursi	Ph.D Visiting Fellow	1995
Chun Hong	MD. Ph.D NRSA Fellow	1994
Jing Lou	MD. Ph.D NRS associate	1994
Saverio Minucci	MD Visiting Associate	1992
Kazushige Sasaki	MD Visiting Fellow	1995
I-Ming Wang	Ph.D IRTA	1993
Other		
Angela Thornton	BS Pre-IRTA	1992
Yaffa Rubinstein	BS Pre-IRTA	1995
Ishwar Chandra-		
Mouliswaran	M.Sc Guest worker	1995
Shawn Kahn	Professional service C	1995
Karyn Cheng	Summer IRTA	
Veronique Nussenblatt	Summer IRTA	
Tom Narangue	Summer IRTA	

Total Man year 11.2 Professional 9.5 Others 1.7

Objectives

The overall objective of this program is to elucidate the mechanisms of gene regulation relevant to the development of the immune system. Within this framework, we analyze transcription of MHC class I and other genes that are regulated by retinoids and interferons.

We have previously isolated and studied several genes encoding proteins that bind to the regulatory DNA elements of MHC class I and other genes. These genes are regulatory genes that belong to large families, and encode proteins that contain distinct DNA binding domains. We found that these DNA binding proteins bind cis elements of not only MHC class I genes but many other genes. Based on the multiple binding specificities we have proposed a "combinatorial model" for gene regulation. In this model regulation of individual genes is controlled by multiple transcription factors bound to multiple regulatory elements, and the patterns of the combination of factors determine specificity of transcription. This is because transcription factors are not unique to each gene but are shared by many genes. From this model it is expected that the functions of these DNA binding proteins are broad, and pleiotropic. The main objective of our research is to study their regulatory roles.

To analyze the mechanisms of transcription, it is important to establish a model system in which each and every component involved in transcription can be identified and dissected. An in vitro transcription assay using the G-free cassette is suitable for this purpose. Our current goal is to apply this assay for studying mechanisms of the action of cloned DNA binding proteins.

In the living cells chromosomal genes interact with nucleosomes and chromatin and are organized into higher orders. In the intact nucleus, cis regulatory elements may be exposed to a special microenvironment, where some transcription factors are locally concentrated or depleted. Such microenvironments may be far from the environment in nuclear extracts prepared in a tube, in which we study DNA-protein binding. Thus gene regulation studied in vitro offer only circumstantial evidence, until it is proven in vivo. To directly demonstrate binding of nuclear factors to the cis elements of genomic genes in the nucleus, in vivo DNA foot print analysis is pursued, which can be achieved by using the PCR mediated DNA foot print method.

Methods employed

Transfection assays.

To study functions of RXR β and ICSBP we use primarily co-transfection assays. Expressible cDNAs are constructed by placing a promoter sequence in front of full length cDNA or dominant negative constructs. The promoter region of the human β -actin gene, RSV LTR or CMV enhancer (pcxn) is used to drive expression of the cloned cDNA. These expression vectors are transfected into

fibroblasts, embryonal carcinoma cells, lymphocytes, and macrophages together with a test reporter gene. The reporter constructs are constructed with natural as well as artificially constructed promoters driving the fire fly luciferase gene. Various methods of transfection including the calcium phosphate precipitation method, DEAE dextran method, electroporation, lipofection are investigated for maximizing the efficiency of transfection. Following incubation for 16 to 48 hrs, cells are harvested and subjected to luciferase analysis. Transfection efficiency is normalized by cotransfected β -galactosidase activity.

In vitro transcription

In vitro transcription assays using G-free cassettes are extensively used to study how RXR β and its heterodimer partner activate ligand dependent transcription. Thyroid hormone responsive elements such as DR-4 are connected to a 40 bp long basal promoter consisted of the TATA element and the initiation site driving a G-free cassette. Another template with a shorter G-free cassette driven by the same basal promoter without responsive elements is used as a control template. As a source of recombinant receptors, extracts prepared from recombinant RXR β or T3R are used. The receptors and templates are mixed with nuclear extracts from B lymphocytes (containing basal transcription factors), NTPs and ligands (T3 or 9cis RA). Levels of transcription is measured by run off assays.

Genomic footprinting.

To study promoter occupancy of a genomic gene ligation mediated PCR has been extensively used in this laboratory. Cells are treated with dimethyl sulfate for 2 min and high molecular weight DNA isolated and digested with piperdine. By using an oligonucleotide primer corresponding to part of the gene the promoter sequence is extended to various positions. The extended fragments are blunt-end ligated by a common primer. This allows exponential amplification of a promoter according to methylation. The amplified fragments are resolved in a sequencing gel.

Expression

To assess a function of the regulatory genes we have isolated, it is necessary to study their expression at the level of protein. To this end monoclonal and polyclonal antibodies are raised against recombinant proteins produced in the baculovirus or in bacterial expression vectors. Mice are immunized with crude lysates containing recombinant proteins, and hybridomas producing specific antibodies are screened by positive and negative selection. Antibodies from the established hybridomas are tested for immunoreactivity with cellular proteins by using Western blot analysis. With these antibodies intracellular localization, tissue and developmental specificity of the expression of those immunoreactive proteins are studied. DNA binding activity of the immunoreactive materials will be tested with super mobility shift, DNA-protein blot assay and immunoprecipitation procedures.

Major findings

Vesicular Stomatitis Virus Infection Induces A Nuclear DNA Binding Factor Specific for the Interferon Stimulated Response Element (primary contributions by Chiara Bovolenta, Jing Lou, Yuka Kanno)

Vesicular stomatitis virus (VSV) has a broad host range. It replicates in the cytoplasm and causes rapid cytopathic effects. We show that following VSV infection, a nuclear factor that binds to a select set of the interferon stimulated responsive element (ISRE) is induced in many cell types. This factor, designated VSV-induced binding protein (VIBP), was estimated to have an approximate Mr. of 50 kDa, and was distinct from known members of the IFN regulatory factor (IRF) family, that are known to bind to the ISRE. Induction of VIBP required tyrosine kinase activity, but did not require cellular transcription. Treatment of cells with cycloheximide that inhibits translation, only partially inhibited induction of VIBP. However, type I interferons and staurosporine, both of which inhibit VSV transcription, inhibited VIBP induction. Moreover, a double stranded RNA analogue, poly(I)poly(C) also induced a DNA binding activity very similar to that of VIBP. These results indicate that a pre-existing cellular protein is activated upon VSV infection and that this activation requires primary viral transcripts. The functional activity of VIBP was analyzed in cells stably transfected with a tk-luciferase reporter gene that is under control of the ISRE. While activity of the control promoter without ISRE was strongly inhibited following VSV infection (due to virus-mediated transcriptional shutdown of the host cell), the inhibition was reversed by the ISRE containing promoter, albeit partially, which suggests that VSV infection differentially affects transcription of host genes. Although VIBP was induced in all other cells tested, it was not induced in embryonal carcinoma cells after VSV infection, suggesting developmental regulation of VIBP inducibility.

TFIIB and Vitamin D Receptor Cooperatively Activate Ligand Dependent Transcription. (primary contribution by Jorge Blanco and I-Ming Wang).

Vitamin D ($1,25\text{ (OH)}_2\text{D}_3$) regulates gene transcription through binding to the vitamin D receptor (VDR), a member of the nuclear hormone receptor superfamily. Sequence specific transcription factors, including nuclear hormone receptors, are thought to regulate transcription by interacting with the basal transcription complex. In GST based protein-protein binding assays we found that VDR specifically binds to TFIIB, a component of the basal complex, and the interaction requires specific domains of each protein.

To assess functional significance of the interaction, transfection assays were performed with a vitamin D responsive reporter. In P19 embryonal carcinoma (EC) cells co-transfection of VDR and TFIIB cooperatively activated reporter transcription, while each factor alone gave very low to no activation. This activation was dependent on $1,25\text{ (OH)}_2\text{D}_3$, and the dose of TFIIB and VDR transfected. These results demonstrate that a nuclear hormone receptor and TFIIB functionally interact with each other, leading to activation of ligand dependent transcription in vivo. In NIH 3T3 cells, however,

strong reporter activation was observed by VDR alone, and co-transfection of TFIIB led to specific, and dose dependent repression of reporter activity. Taken together, our results suggest that TFIIB-nuclear hormone receptor interaction plays a critical role in ligand dependent transcription, which involves a cell type specific accessory factor.

Inhibition of Ligand Induced Promoter Occupancy in vivo

by a Repressor RXR β (contribution by Jorge Blanco, Anup Dey, and Saverio Minucci). The region of RXR β comprising the C-terminal 21 amino acids is highly conserved in the nuclear receptor superfamily and acts as an autonomous activation domain. We show that deletion of these amino acids converts RXR β into a repressor (designated Δ C2), that silences vitamin D and retinoic acid (RA) dependent transcription, even though Δ C2 retained the ability to heterodimerize with partner receptors and bound to the cognate responsive element in vitro. P19 embryonal carcinoma (EC) cell clones that stably express Δ C2 were established. These clones were deficient in activating an RA responsive reporter, and in early induction of the endogenous RAR β gene upon RA treatment, both of which depend on activity of RXR/RAR heterodimers. To elucidate the mechanism of Δ C2 action we studied the occupancy of the RAR β_2 promoter in these clones by genomic footprinting analysis. The RAR β_2 promoter contains a canonical RA responsive element (RARE, to which unliganded RXR/RAR heterodimers bind in vitro) and several other elements, all of which were not occupied prior to RA treatment in P19 cells. In control clones, prominent protection was induced at the RARE and all other elements after RA treatment. In clones expressing Δ C2, in contrast, protection was markedly inhibited at the RARE as well as all other elements in the promoter. Thus, the defect in ligand induced promoter occupancy in vivo is likely to account for the silencing activity elicited by Δ C2. Taken together, these results indicate that the C-terminal domain of RXR β is critically involved in ligand induced occupancy of the responsive element in vivo, an event that leads to the recruitment of other factors to their cognate elements in the promoter.

Inhibition of HIV-1 and Vaccinia Virus Infection

by a Dominant Negative Factor of the IRF Family Expressed in Monocytic Cells (primary contribution by Angela M. Thornton I-Ming Wang).

ICSBP is a member of the interferon regulatory factor (IRF) family that regulates expression of type I interferon (IFN) and IFN regulated genes. To study the role of the IRF family in viral infection, a cDNA for the DNA binding domain (DBD) of ICSBP was stably transfected into U937 human monocytic cells. Clones that expressed DBD exhibited a dominant negative phenotype and did not show anti-viral activity against vesicular stomatitis virus (VSV) infection in response to IFN. Furthermore, vaccinia virus (VV) and human immunodeficiency virus-1 (HIV-1) failed to infect clones expressing DBD. The inhibition of VV infection was attributed to defective virion assembly, and that of HIV-1 to low CD4 expression and inhibition of viral transcription. HIV-1 and VV were found to

have sequences in their regulatory regions, similar to the IFN stimulated responsive element (ISRE), to which IRF family proteins bind. In accordance, these viral sequences and a cellular ISRE bound a shared factor(s) expressed in U937 cells. These observations suggest a novel host-virus relationship in which the productive infection of some viruses is regulated by the IRF dependent transcription pathway through the ISRE.

Expression of Interferon Regulatory Factor (IRF) Family Proteins in Lymphocytes. Induction of Stat 1 and ICSBP Expression by T cell Activation (primary contribution by Nancy Nelson and Yuka Kanno). ICSBP is a transcription factor belonging to the interferon regulatory factor (IRF) family that regulates IFN mediated gene expression. Several lines of work indicate that the IRF family plays an important role in the immune system. Although some proteins of the IRF family are expressed in many types of cells, ICSBP is exclusively expressed in the immune system. In this work immunoblot analysis was performed to study expression of ICSBP and other members of the family in various murine lymphocytes. The results show that ICSBP was expressed constitutively in B cells throughout development, and in resting and activated cells. In contrast, ICSBP expression was undetectable in thymocytes and resting T cells, while all other IRF proteins tested (IRF-1, IRF-2 and ISGF3 γ) were detected in these cells. However, strong ICSBP induction (and weak IRF-1 induction) was observed upon mitogenic activation of T cells. Expression of other IRF proteins was unaffected by T cell activation. Once T cells were activated, ICSBP was stably expressed in these cells including TH1 and TH2 cells. We found that mitogenic activation of T cells leads to induction of Stat1 that binds to the IFN- γ responsive element in the ICSBP and IRF-1 promoters, which most likely accounts for their induction. Stat1 induction (and hence ICSBP and IRF-1 induction) in mitogen-activated T cells was mediated by IFN- γ , since T cells from IFN- γ ^{-/-} mice failed to induce both Stat1 and ICSBP. However, anti-CD3 Ab binding to T cell hybridoma also led to induction of Stat1 and ICSBP, even though IFN- γ was not produced by these cells. Taken together, these results demonstrate that the function of ICSBP is coupled with T cell activation, which, in some cases, involves a novel pathway of Stat1 activation, independently of IFN γ stimulation.

Phenotypic analysis of ICSBP^{-/-} mice (primary contribution by Yuka Kanno and Jing Lou).

In collaboration with Dr. I. Horak, the ICSBP gene in mouse chromosome 6 has been disrupted by homologous recombination, and mice with germ line transmission have been established. Mice homozygous with ICSBP^{-/-} are now being investigated. We have confirmed that these mice lack expression of ICSBP proteins in lymphoid cells and lymphoid cells are deficient in factor binding activity specific for the ISRE. The ICSBP^{-/-} mice show a striking phenotype in that the granulocyte population (as defined by expression of the 8C5 marker) is increased more than 10 fold in bone marrow, spleen, lymphonodes and in peripheral blood. In

addition, the macrophage population (with the Mac1 marker) is increased by >5 fold in these organs, although T cells and B cells are roughly normal in these cells in terms of their number and surface property. Although ICSBP^{-/-} mice have immunoglobulins with all isotypes, the levels of IgE and IgG1 are abnormal and are 10 fold greater than that of normal counterparts. Levels of various cytokines IL-2, IL-4, IL-6, IL-10, IFN- γ , IFN α , IFN β , TNF α appear normal in ICSBP^{-/-} mice. Production of IFN γ upon LPS treatment, however, appears to be defective in the knock-out mice. Interestingly, these mice are highly susceptible to infection by parasite *Listeria* and mouse Herpes virus.

Proposed course of work

The mechanism of TFIIB-factor interaction.

In vitro approach.

To identify molecules involved in forming a complex with TFIIB, the following in vitro assay will be used and by which the mode of interaction will be biochemically analyzed. Firstly, biotinylated oligonucleotides for the RARE or the ISRE will be immobilized to the magnetic beads, and recombinant, purified RXR β and RAR β will be bound to the beads through the specific RARE site. Nuclear extracts from P19 EC cells will be fractionated on a P11 column that contain different basal transcriptional activities obtained from P19 EC cells will be added to the complex in the presence or absence of TFIIB, and retinoic acid. The complex formed will be analyzed with silver staining and western blot analysis. To ensure the specificity of binding, various mutant receptors including δ C2 will be used. To study the functional activity of the complex, materials will be eluted from the beads and tested in an in vitro transcription assays. As an alternative approach, in vitro transcription templates immobilized onto beads may be used to perform above experiments and transcription directly measured. Based on our previous functional data demonstrating cell type specificity, complexes formed this way are expected to contain cell type specific factors. Upon identification of specific bands we will isolate the peptides. These peptides will be further analyzed biochemically for its charge and phosphorylation properties. Peptides will be microsequenced through a collaborative arrangement, and processed for cloning. These experiments will also be performed using IRF-1 and IRF-2 instead of RAR, and RXR.

In vivo approach

An important question that can be addressed in vivo is the role of TFIID and the initiator sequence in TFIIB-factor interaction. Various mutant initiator reporters will be constructed and tested for their activity in supporting TFIIB- factor interactions in P19 and NIH 3T3 cells. We will also construct mutant TBP and TFIIB in a mammalian expression vector and examine their ability to mount a functional interaction in various cells. The role of USF and YY1 will be addressed as both of them are shown to be involved in initiator activity.

Analysis of ICSBP^{-/-} mice:

Viral infectivity. Mice lacking the expression of ICSBP will be tested for their defense against viral infection. The following viruses will be examined: mice will be infected with murine influenza, murine herpes virus, VSV, MAIDS, and MuLV. Viral clearance, morbidity of infected mice and production of IFNs and cytokines will be measured. Upon initial work with the whole animal we will infect isolated lymphocytes and macrophage populations with the respective viruses in vitro and study molecular events focusing on transcription of ISRE containing promoters. Collaborations with several groups familiar with the

above viruses are being established. Analysis of ICSBP^{-/-} mice so far indicated that expression of IRF-2 is strongly repressed in these mice. We will study the basis of the repression by promoter analysis, western blot assays, and EMSA. IRF-2 and ICSBP share a similar function in that they repress ISRE promoter activity. It will be of interest to study mice that are deficient in both IRF-1 and ICSBP. IRF-2^{-/-} mice produced by Mak's group is available on a collaborative basis. To Construct double knock-outs We will cross IRF-2^{-/-} mice with ICSBP^{-/-} mice to produce double knock-outs, which will then be tested for a phenotype and viral infectivity as above.

Significance to Biomedical Research and the Program of the Institute

One of the central questions of modern biology is the mechanisms of gene regulation. This question is also the fundamental theme of developmental biology. Our program addresses the molecular basis involved in the development of the immune system. Isolation and the subsequent studies of the regulatory genes we undertook in the past year have provided not only new knowledge but also new concepts for gene regulation, which are relevant to basic biology as well as to understanding of clinical problems. Investigation of new means for dissecting the mechanisms of gene regulation pursued in this program is also consistent with overall mission of the Institute.

Publications:

Journal Articles:

Becker K, Jedlicka P, Templeton N, Liotta L, Ozato K. HUCRBP (YY1,NF-E1, δ) is a transcription factor that binds the regulatory regions of many viral and cellular genes, *Gene* 1994;150:259-66.

Becker KG, Nagle JW, Canning RD, Biddison WE, Ozato K, Drew PD. Signature sequencing: Rapid isolation and characterization of 118 novel C2H2-type zinc finger cDNAs expressed in human brain, *Hum Mol Genet* 1995;4:685-91.

Blanco J, Wang I-M, Tsai S, O'Malley B, Jurutka P, Haussler M, Ozato K. TFIIB and the vitamin D receptor cooperatively activate ligand dependent transcription, *Proc Natl Acad Sci USA* 1995;92:1535-9.

Bovolenta C, Lou J, Kanno Y, Park B-K, Thornton A, Coligan J, Schubert M, Ozato K. Vesicular stomatitis virus infection induces a nuclear DNA binding factor specific for the interferon stimulated response element, *J Virol* 1995;69:4173-81.

Dey A, Minucci S, Ozato K. Ligand dependent occupancy of the RAR β 2 promoter in vivo, *Mol Cell Biol* 1994;14:8191-8201.

Drew P, Franzoso Q, Becker K, Bours V, Carlson L, Siebenlist U, Ozato K. NFkB and Interferon Regulatory Factor-1 physically interact and synergistically induce major histocompatibility class I gene expression, J Interferon and Cytokine Res 1995;In Press.

Drew PD, Nagle JW, Canning RD, Ozato K, Biddison, WE, Becker KG. Cloning and expression analysis of a human cDNA homologous to *Xenopus TFIIIA*, Gene 1995;159:215-8.

Gu J, Harriss J, Ozato K, Gottlieb P. Induction by Con A of specific mRNAs and cytolytic function in a CD8-positive T-Cell hybridoma, J Immunol 1994;153:4408-17.

Jungwirth C, Rebbert M, Ozato K, Degan HJ, Schultz U, Dawid IB. Chicken ISCSF and IRF-1 genes reveal evolutionary conservation in the IRF gene family, Proc Natl Acad Sci USA 1995;92:3105-9.

Kalvakolanu, DV, Mannino, SB, Thornton A, Ozato K, Borden EC. Transcriptional induction of genes by IFN- β in mouse cells is regulated by a transcription factor similar to human ISGF-3, Antiviral Res 1995;In Press.

Leng X, Blanco J, Tsai SY, Ozato K, O'Malley BW, Tsai M-J Mechanisms for synergistic activation of thyroid hormone receptor and retinoid X receptor on different response elements, J Biol Chem 1994;269:31436-42.

Leng X, Blanco J, Tsai SY, Ozato K, O'Malley BW, Tsai, M-J Mouse retinoic x receptor contains a separable ligand binding and transactivation domain in its E region, Mol Cell Biol 1995;15:255-63.

Martinez-Balbás MA, Dey A, Rabindran SK, Ozato K, Wu C. Displacement of sequence-specific transcription factors from mitotic chromatin, Cell 1995;In Press.

Medin, JA, Minucci S, Driggers PH, Lee I, Grippo JF, Ozato K. Quantitative increases in DNA binding affinity and positional effects determine 9-cis retinoic acid induced activation of the retinoid X receptor beta homodimer, Mol Cell Endocrin 1994;105:27-35.

Nunez S, Medin J, Keller H, Wang K, Ozato K, Wahli W, Segars J. Retinoid X receptor β and peroxisome proliferator-activated receptor activate an estrogen response element, Recent Prog Hormone Res 1994;In Press.

Nagata T, Weiss E, Abe K, Kitagawa K, Ando A, Yara-Kikuti Y, Seldin M, Ozato K, Inoko H, Taketo M. Molecular mapping of the retinoid

X receptor β gene in mouse and man, Immunogenetics 1995;41:83-90.

Yang Y, Minucci S, Zand D, Ozato K, Ashwell J. T cell activation and increases in protein kinase C activity enhance retinoic acid-induced gene transcription, Mol Endocrinol 1994;8:1370-6.

BOOKS:

Samuel CE, Ozato K. Induction of interferon and interferon-induced genes. In: Pestka S, Schellekens H, eds. Cytokine Yearbook, Biotherapy. Dordrecht, the Netherlands: Kluwer Academic Publisher, 1995;In Press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 08719-15 LMGR

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Development and Uses of Eukaryotic Vectors

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute)

P.I.:	Bruce H. Howard	Chief	LMGR, NICHD
Other:	M.L. Avantaggiati	Visiting Fellow	LMGR, NICHD
	E. Brinkmann	Technician	LMGR, NICHD
	V. Ogryzko	Visiting Fellow	LMGR, NICHD
	E. Englander	Visiting Associate	LMGR, NICHD
	G. Humphrey	Visiting Associate	LMGR, NICHD
	V. Russanova	Visiting Fellow	LMGR, NICHD
	K. Engleka	NRC Bio Tech	LMGR, NICHD

COOPERATING UNITS (if any)

LAB/BRANCH

Laboratory of Molecular Growth Regulation

SECTION

Section on Human Cell Genetics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892-2753

TOTAL STAFF YEARS:

7.50

PROFESSIONAL:

6.50

OTHER:

1

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Cell cycle checkpoint mechanisms in mammalian cells act through inhibition of cyclin/cyclin-dependent kinase complexes, as well as through tumor suppressor gene products belonging to the retinoblastoma (Rb) gene family. Work over the past year has demonstrated that these anti-proliferative mechanisms influence multiple phases of the cell cycle. For example, the p21^{WAF1/CIP1/sdi1} gene product retards S phase progression in a manner that is antagonized by cyclin A, cyclin E, or the phosphatase 2A-binding activity of simian virus 40 small t antigen. These findings lead to the conclusion that cyclin-dependent kinase activity is involved in the regulation of S phase progression in mammalian cells. While activation of this and other anti-proliferative pathways by DNA damage, serum deprivation, etc. has been intensively investigated, there is little information on how such pathways are triggered during the onset of cellular senescence. Experiments in which histone deacetylase inhibitors were shown to accelerate the onset of a senescence-like state point to one possible explanation, namely, that defects in higher order chromatin structures can activate checkpoint mechanisms, thereby leading to irreversible cell cycle exit.

Project DescriptionObjectives:

The Human Genetics Section investigates molecular mechanisms by which the mammalian cell cycle is controlled. Using molecular cloning and gene transfer technologies in combination, the research program of this section focuses on pathways mediating quiescence, terminal differentiation, and senescence in normal human cells. Attention is given to well characterized effectors of cell cycle regulation such as cyclins, cyclin-dependent kinases and their inhibitors, as well as tumor suppressor genes of the *retinoblastoma* (*Rb*) family. Considerable effort is devoted, in addition, to less well established growth regulatory mechanisms that involve epigenetic gene regulation. Here the possibility is investigated that higher order repressive chromatin structures in cells are monitored, and that defects in the assembly of such structures can activate cell cycle checkpoints, thereby causing growth arrest.

Major Findings:

A major project in the **Human Genetics Section** concerns mechanisms underlying cellular immortalization and its converse, cellular senescence. Cellular senescence in human diploid fibroblasts (HDF) and other human cell types has been widely studied, yet in many aspects remains poorly understood. One area which in particular merits further investigation is the apparent relatedness of cellular senescence and differentiation. Although the two processes are not necessarily synonymous, in many respects senescence resembles a *forme fruste* of differentiation, with cells acquiring a phenotype that is suboptimal for tissue function and maintenance. HDF proliferative potential is primarily dependent on the number of rounds of DNA synthesis completed rather than the cumulative time in culture; thus, to the extent that senescence is considered as akin to terminal differentiation, it most closely resembles model systems (e.g., murine erythroleukemia or promyelocytic leukemia cells) where a requirement for passage through the cell cycle has been demonstrated.

Our attention to the differentiative aspects of senescence was first prompted by experiments designed to compare cell cycle arrest mechanisms in senescent and quiescent cells. Transient expression of simian virus 40 (SV40) T antigen is strongly mitogenic, i.e., is dominant over cell cycle arrest mechanisms, in both senescent cells and serum-deprived quiescent fibroblasts. We and others observed that in senescent HDF, but not quiescent cells, the domain in T antigen which mediates binding to *Rb* family proteins is required for efficient stimulation of DNA synthesis. This requirement is likely to reflect the failure in senescent cells to downregulate p21^{Waf1/Cip1/Sdi1}, which in turn leads to accumulation of the *Rb* gene product (pRb) in its anti-proliferative hypophosphorylated form. In the course of these and related experiments, it was noted that HDF deprived of serum in the presence of either of two histone deacetylase (HD) inhibitors, sodium butyrate or trichostatin A, resemble senescent HDF by the above criteria; that is, both senescent and HD-treated cells exhibit a response to the mitogenic action of T antigen that is dramatically attenuated when the T antigen pRb-binding domain is inactive. These results led us to consider the possibility that the antiproliferative pathways activated by HDs might be similar to those operative in senescent HDF. If treatment with such agents could be shown to hasten entry of HDF into a senescence-like state, we reasoned, it would reinforce the view that a major component of HDF senescence is mechanistically related to differentiation, particularly since cell cycle-dependent terminal differentiation in several experimental systems has been shown to be inducible by HDs.

Over the past year, a number of experiments have supported the notion that HDs can indeed reduce proliferative lifespan in human fibroblasts, causing those cells to enter a senescence-like state prematurely. As predicted from the above hypothesis, cells propagated in butyrate or

trichostatin A entered a G1 phase arrest state after several cell doublings. When such prematurely arrested cells (PA cells) were released from these agents, they remained in G0/G1 arrest and could be shown to resemble senescent cells by the following criteria: i) expression of senescence biomarkers (senescent and PA cells were alike in terms of cell size, saturation density, and expression of pH 6 β -galactosidase, a recently described marker that distinguishes senescent from quiescent cells); ii) hypophosphorylation and underexpression of pRb; iii) competence to initiate a round of replicative DNA synthesis in response to SV40 T antigen; and iv) attenuation of mitogenic responsiveness to a T antigen mutant defective for binding to the retinoblastoma gene product, pRb. Notably, the induction of a senescence-like state by HDs was found to be dependent on one or more cell doublings in the presence of these agents, consistent with the view that they act on the senescence-associated cell cycle counting mechanism.

The immediate question that arises from the above findings is how histone deacetylase inhibitors might act to reduce proliferative potential. There are two interrelated possibilities to be considered, the first being that these agents act locally to modulate the expression of cell cycle-regulatory genes. Hyperacetylation of histones H3 and H4 correlates with actively transcribed regions, and butyrate has been shown to facilitate enhancer function. On the other hand, butyrate treatment can either induce or block induction of gene expression, depending on the target gene studied. Importantly, histone hyperacetylation is reversible after butyrate is removed, and, where examined, transcription of butyrate-induced genes returns to its basal level after the cessation of treatment. The simplest way to reconcile the reversibility of histone deacetylase inhibitors (i.e., on acetylation and local transcription) with the apparently irreversible cell cycle exit that these agents can induce is to postulate that cell doubling in the presence of these agents alters a second, as yet undefined, control mechanism. DNA methylation, for example, is a candidate for such a coupled mechanism; thus, DNA replication in the presence of butyrate or trichostatin could conceivably induce irreversible loss of DNA methylation at critical regulatory loci. Effects of butyrate on the level of 5-methylCpG modification have been reported; however, the patterns of change are complex, including both increased and decreased DNA methylation.

The second possibility to be considered is that butyrate and trichostatin A act indirectly by affecting the stability of relatively large repressive chromatin structures. Such structures, which extend over a few kilobase pairs to several tens of kilobase pairs (or more), are thought to mediate epigenetically inherited silencing, are implicated in the developmental control of differentiation, and in specific cases appear to persist by virtue of memory mechanism(s). A crucial feature of heterochromatin-like regions, indeed one which is conserved from yeast to humans, is the underacetylation of selected lysine residues in histone H4; accordingly, proper function of histone deacetylase is assumed to be required for the reformation of these regions during the latter part of each cell cycle. While experiments directly pertaining to heterochromatin structure are currently in progress, it has been shown in *Drosophila* that sodium butyrate acts as a suppressor of position effect variegation, and that this agent enhances phenotypic effects of certain suppressor of position effect variegation mutants. Since suppression of position effect variegation is generally believed to reflect destabilization of heterochromatin domains, such results strongly suggest that butyrate affects those structures. Coupling the foregoing with the idea that a critical fraction of heterochromatin-like regions in non-immortalized HDF may persist by memory, it is tempting to speculate that manipulations which disrupt heterochromatin formation during late S and/or G2 phases in those cells should predispose to irreversible cell cycle exit.

The suggestion that heterochromatin-like domains may play a role in aging is not without precedent. Correlative evidence consistent with selective heterochromatin instability as a contributory factor in senescence derives from reports of age-dependent recession of a subset of X-chromosome associated heterochromatin domains in vivo and in vitro. Moreover, a recent study on aging in *Saccharomyces cerevisiae* independently reached the conclusion that the stability of repressive chromatin domains may be a strong determinant of proliferative potential. The latter study does not directly address the issue of histone acetylation in chromatin-mediated silencing, but

elegant investigations on regulation of the mating-type loci *HML* and *HMR* point to a causal role for this histone modification in the maintenance of repressive chromatin structures at those sites. To the extent that our results and those obtained in the yeast system are in agreement, they complement each other and greatly strengthen the argument that *Saccharomyces cerevisiae* provides an informative model for studies on mammalian cell senescence. Further progress in this area should be realized by a combination of genetic analysis in yeast, and, in mammalian cells, manipulation of genes involved in histone acetylation and heterochromatin metabolism.

A second major project in this section concerns the roles of cyclins, cyclin-dependent kinases, and associated tumor suppressor gene products. The latter include pRb and a newly discovered protein, CAP, which is described in more detail in the section pertaining to Dr. Yoshihiro Nakatani's research program. Interest in this area stems from an appreciation that cell cycle checkpoints, if indeed activated in some situations by defects in higher order chromatin structures, must ultimately act through downstream effector molecules. Inhibitors of cyclin-dependent kinases such as the p21^{WAF1/CIP1/sdi1} gene product (WAF1) are particularly good candidates to serve in this role.

Over the past year, it was demonstrated that WAF1 not only blocks S phase entry by inhibition of cyclin-dependent kinases (CDKs), but also can also impede S phase progression. The latter WAF1-mediated inhibitory effect could be antagonized by cyclin A, cyclin E, or the phosphatase 2A-binding activity of simian virus 40 small t antigen. From these findings we were able to conclude for the first time that CDK activity is involved in the regulation of S phase progression in mammalian cells. This conclusion is in agreement with genetic data from yeast and *in vitro* data from *Xenopus* oocytes.

A related issue that was addressed is whether the primary target for WAF1-mediated intra-S phase arrest is PCNA or cyclin/CDK complexes. DNA replication *in vitro* can be directly inhibited by WAF1 through the PCNA binding, and it has been suggested that S phase progression *in vivo* can be inhibited by this mechanism. Evidence obtained in this work, however, implicated CDKs as the most likely targets of WAF1 induction. The reason for this may be a relative abundance of PCNA compared to components of the CDK machinery, as has been suggested for the *Xenopus* oocyte system; accordingly, relatively high levels of WAF1 expression might be necessary to observe a direct block of DNA replication due to PCNA titration.

Publications

Avantaggiati ML, Carbone M, Graessman A, Howard BH, Levine AS. The SV40 large T antigen and adenovirus E1a oncoproteins interact with different isoforms of the transcriptional coactivator, E M B O J 1 9 9 5 ; i n p r e s s .

Englander EW, Howard BH. Nucleosome positioning by human Alu elements in chromatin. J Biol C h e m 1 9 9 5 ; 2 7 0 (1 7) : 1 0 0 9 1 - 9 6 .

Howard BH, Russanova VR, Englander EW. Alu silencing mechanisms: implications for the modulation of local chromatin structure. In: Maraia RJ, ed. Impact of short interspersed elements (SINEs) on the host genome. Austin: RG Landes Co, 1995; 133-41.

Howard BH. Replicative senescence: considerations relating to the stability of heterochromatin domain, Exp Gerontology 1995;31(1/2):in press.

Jazwinski SM, Howard BH, Nayak RK. Cell cycle progression, aging, and cell death, J Gerontology Biol Sci 1995;50A:B1-B8.

Ogryzko VV, Howard BH. Is cell cycle progression coupled to the maintenance of heterochromatin-like domains, TIBS 1995;in press.

Russanova VR, Driscoll CT, Howard BH. Adenovirus type 2 preferentially stimulates polymerase III transcription of Alu elements by relieving repression: a potential role for chromatin, Mol Cell Biol 1995; 15 (8) : 4282 - 90 .

Vorce RL, Lee B, Howard BH. Methylation- and mutation-dependent stimulation of Alu transcription in vitro, Biochem Biophys Res Comm 1994;203:845-51.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 00165-20 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Isolation and Characterization of Macromolecular and Cellular Particles

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: A. Chrambach Head

Others: A. Sokoloff Visiting Associate
S. Radko Courtesy Associate
H. Chang Visiting Fellow
G. Caiafa

COOPERATING UNITS (if any)

None

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Macromolecular Analysis

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

0.5

PROFESSIONAL:

0.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
- ☐ (a1) Minors
- ☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Some of the microsome components formed under conditions of microsome fusion were previously detected in capillary zone electrophoresis using buffered polyacrylamide solutions. We attempted to isolate these components by 1) modification of design and procedure of free-flow electrophoresis, and 2) by application of automated gel electrophoresis apparatus using agarose gels and agarose or polyacrylamide solutions. Both attempts have failed to date but are being pursued further.

LTPB-FY95



Project Description

The project aims at applying advances in separation technology reported in the joint Project Report to macromolecular and cellular particle separations of biological importance. In particular, that technology comprises quantitative and automated modes of electrophoresis in gels and polymer solutions, and associated instrumentation and software.

Objective:

This year's objective was the isolation and subsequent characterization of some microsomal cell wall components which are produced during microsome fusion. The past year had produced evidence of a number of such components in capillary zone electrophoresis, using polyacrylamide (of molecular weight 5 million) solutions as the "molecular sieving" medium. Since the pattern of peaks was too complex to allow for the construction of Ferguson plots of these components (in view of the fact that mutual overtaking of bands upon variation of polymer concentration could not be ruled out), isolation became necessary for Ferguson plot analysis. Capillary zone electrophoresis does not lend itself to isolation in the required amounts. Thus two approaches were attempted which promised to yield the required isolation method - automated apparatus using agarose gels and solutions, and a modified Hannig apparatus in conjunction with polymer solutions.

Subproject (1): Free-flow electrophoretic isolation of microsome components in buffered polymer solutions (Radko, S. P., Sokoloff, A. and Chrambach, A.)

Objective:

Isolation by free-flow electrophoresis in buffered polymer solution of microsome components arising under conditions of microsome fusion.

Major Findings:

Free flow electrophoresis (FFE) apparatus was modified to provide as close as possible capillary dimensions in the separation chamber - 0.3 mm in lieu of the 0.2 mm maximally used in capillary zone electrophoresis. Mimicking the conditions of buffered polyacrylamide solution that had been effective in microsome separations, polystyrene size standards in the size range of the microsome gave rise to two or more components in free flow electrophoresis while under the identical buffer and polymer conditions they appeared homogeneous upon capillary zone electrophoresis. The viscosity of polymer solutions upon continuous flow into the separation chamber apparently gives rise to this artifact which we are still in the process of remedying before microsome products can be applied meaningfully.

Proposed Course:

The artifactual bands in FFE arise only in presence of polymers. By using polymer concentrations and different types of polymer varying in viscosity we will test whether the flow characteristics giving rise to parabolic bands can be regulated by viscosity to suppress the formation of artifacts. An attempt will be to polymer-coat the plastic wall of the separation chamber, in addition to covalent coating of the glass wall, to suppress electroendosmosis and its effect producing parabolic zones. A technique of interrupting flow while electrophoresis proceeds will attempt to achieve a uniform separation path extending over the entire width of the separation chamber. Phycoerythrin will be used to define the length of the separation path and relative mobility values for the characterization of the microsome components.

Significance to Biomedical Research and the Program of the ICD:

Identification and physical characterization of the microsomal components participating in fusion will promote our understanding of the mechanism of membrane fusion.

Subproject (2) - Isolation of microsome components on automated gel electrophoresis apparatus using either gels or polymer solutions (Chang, H.-T., Zakharov, S. F., Caiafa, G. and Chrambach, A.)

Objective:

Isolation of some of the microsome components arising under conditions of fusion and observed but uncharacterizable in capillary zone electrophoresis serves the purpose of their identification and physical characterization.

Major Findings:

Both agarose gels and agarose solutions in the concentrations range 0.05 to 3.0% adsorbed microsomes at the position of sample application, presumably due to irreversible entanglement with the agarose fiber. Microsome migration proceeds satisfactorily in polyacrylamide solution, but no separation could as yet be demonstrated in application of any polymer solution in the 3 mm thick separation chamber of the automated apparatus, although polystyrene size standards form bands the dispersion of which decreases with increasing polymer concentration. While separation conditions are being investigated, a method for band isolation from polymer solutions in the automated apparatus has been developed which uses the optical system of the apparatus to position a syringe over the band for its volumetric recovery (subproject 6 of the joint report).

Proposed Course:

Since band dispersion decreases with polymer concentration under the conditions of automated electrophoresis in polymer solutions used, we will determine the maximal polymer concentration at which microsome migration proceeds at a practical rate and test separation on hand of the well-known 2-component pattern of purified microsomes. Once separation conditions in any of the polymer types are established, isolation will utilize the capacity of automated apparatus for the simultaneous monitoring of recovery. An instrumental adaptation for polymer solutions (subproject 6 of the joint report) will be applied which replaces the electrophoretic recovery from gels by a volumetric recovery from solution.

Significance to Biomedical Research and the Program of the ICD:

Isolation from polymer solutions by use of automated gel electrophoresis apparatus and the monitored, quantitative recovery technique associated with it, would benefit the biochemistry of subcellular-sized particles in general and our understanding of the membrane fusion mechanisms in particular.

Publications:

Radko SP, Sokoloff AV, Garner MM, Chrambach A. Capillary electrophoresis of rat liver microsomes in polymer solutions, *Electrophoresis* 1995;16:981-992.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 00171-19 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Electrophoretic Methodology

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: A. Chrambach Head LTPB, NICHD

Others: M. Garner Expert LTPB, NICHD
 S. Radko Courtesy Associate LTPB, NICHD
 S. Zakharov Visiting Fellow LTPB, NICHD
 H. Chang Visiting Fellow LTPB, NICHD
 G. Weiss Collaborator DCRT, NIH
 A. Aldroubi Collaborator NCRR, NIH

COOPERATING UNITS (if any)

DCRT (G.H. Weiss); NCRR (A. Aldroubi)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Macromolecular Analysis

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

4.1

PROFESSIONAL:

4.1

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Previous technology allows for densitometry of bands diffusing after the termination of electrophoresis. It also yields a single band profile per experiment, and not the time course of dispersion except in those rare cases where a photographic setup allowed one to follow the band in time. But even then, the non-linearity of the photographic response made it difficult to interpret band width data. A recently developed automated gel electrophoresis apparatus with intermittent fluorescence scanning of the migration path solved these problems by providing the band profile while electrophoresis was proceeding. Its width at half-height allows for prediction of band dispersion as a function of migration time and therefore of resolution. The measurement of resolution was used 1) to create a computer program specifying the optimally resolving gel concentration and migration time; and 2) to objectively evaluate the resolving power of separation methods and of the "molecular sieving" capacity of polymers. Availability of the band profile was also used to measure deviations from Gaussian band shape. These were compatible with a model of progressive entanglement of the particle in the gel fiber network. The previous theory of dispersion in gels was modified accordingly. The preparative capacity of the automated apparatus was developed by creating a computer program predicting the electroelution time needed for a selected degree of recovery, and by increasing the load volume capacity of the apparatus from 20 ul to 1.5 ml. The apparatus was also applied to electrophoresis of subcellular-sized particles in polymer solutions. A preparative device for that application was constructed. The provision by the automated technique of large numbers of accurate mobility values served to determine their previously unknown precision, thereby validating the basis of quantitative electrophoresis. The theory of "molecular sieving" was further developed by showing that changes in polymer relaxation times, in addition to a threshold dividing semi-dilute and concentrated regimes, accounted for the triphasic relation between retardation and particle or polymer size.

Project Description

The aim of the project is the development of optimally resolving automated methods for the electrophoretic separation of macromolecules and subcellular particles. Since these species differ predominantly in size and shape (conformation), this requires selection of optimally resolving polymer type, concentration and migration time for two components in a particular size range. To be practical, this selection needs to be made by application of computer programs embedding the relevant separation theory and a data base of mobility and band dispersion values. The recent availability of such data from automated apparatus has allowed for the formulation of a first generation of those programs. At this time, theoretical advances are needed to allow for an upgrading of the programs in recognition of non-Gaussian band shapes and electrophoretic dispersion in the absence of diffusion. Of the greatest practical importance is the preparative capacity of electrophoresis in gels and polymer solutions. Automated apparatus allows for the monitoring of recovery steps to attain near-quantitative recoveries from gels. However, data bases for the prediction of recovery times remain rudimentary, separations of subcellular-sized particles in polymer solutions have to be demonstrated and preparative instrument modifications for that purpose still need to mature.

Objective:

This year's work aimed principally at the exploitation of the capacity for band width and shape determination and macromolecular isolations of the automated gel electrophoresis apparatus with intermittent fluorescence scanning of the migration path. Availability of the band profile was to be used to numerically determine resolution, the central parameter of separation science, and to devise a computer program capable of predicting optimally resolving conditions. The measurement of resolution also promised to provide the basis by which electrophoretic separation methods and polymers providing the "molecular sieve" could be selected. Availability of the band profile also promised to test previous notions of diffusion spreading in gels. Finally, the experimental ease provided by automated electrophoresis apparatus promised to furnish the large number and accuracy of mobility measurements needed to statistically evaluate mobility and thereby put quantitative electrophoresis on a firm basis. The preparative application of the apparatus which uniquely allows for obtaining quantitative recovery through a sequential electroelution procedure was to be developed to allow for prediction of electroelution conditions yielding quantitative recovery, and encompass separations of subcellular-sized particles in polymer solutions.

Subproject (1) - Computation of optimally resolving gel (polymer) concentrations and migration times (Aldroubi, A., Chang, H.-T., Zakharov, S. F. and Chrambach, A.)

Objective:

Any pair of components differing in size, shape or conformation is best resolved at a mathematically defined polymer concentration and within a defined migration time. The availability of band profile (width) data from automated gel electrophoresis apparatus has made it possible to calculate optimally resolving conditions under the assumption that the band distribution is Gaussian. The objective of the study was to create a computer program by means of which the optimization of conditions can be carried out practically.

Major Findings:

A computer program was created which employs an input of band width and migration time for each member of a component pair, together with data derived from electrophoretic mobility to predict the optimally resolving gel (polymer) concentration and migration time. The program allows the user to select the degree of distribution (band) overlap for which the prediction is desired. The predicted relation between electroelution field strength and time, and the degree of recovery, was verified experimentally for resolution of fluorescein-labeled β -lactoglobulin from conalbumin.

Proposed Course:

The program needs to be adapted to consider differences in electroelution requirements for bands of different width and for the non-Gaussian band shapes that prevail progressively with increasing gel (polymer) concentration and, in some cases, field strength. The adjustment of the functions relating band dispersion with gel concentration and time, as well as the quantitative expression for deviation from Gaussian band shape developed as described in Subproject 3, will need to be considered. Testing and verification will be extended to the different chemical types of macromolecule.

Significance to Biomedical Research and the Program of the ICD:

The program promises to remedy the entire practice of present-day gel electrophoresis in conducting it at arbitrary gel concentrations and for arbitrary migration times (paths), or those estimated by trial-and-error approaches.

Subproject (2) - A computer program for predicting the optimal polymer type and size for "molecular sieving" in solutions (Aldroubi, A., Radko, S. P. and Chrambach, A.)

Objective:

To replace the present practice of an arbitrary choice of polymer type and size for the purpose of "molecular sieving" in solution by an objective optimization.

Major Findings:

A computer program was created capable of pinpointing, by interpolation of a 4-dimensional "plot", the polymer type and concentration which, for two particles of known size, will provide the maximal separation efficiency (defined by Chrambach, A. and Aldroubi, A. (1993) Electrophoresis 14, 18-22). A data base of the change of retardation with particle size, dKR/dR , for a wide range of particle sizes, polymer types and polymer concentrations provides the input of that program.

Proposed Course:

The predictions of the program still need to be tested and verified widely. Since the data base derives from capillary zone electrophoresis, band width can be evaluated. On its basis, we will modify the program to compute maximal resolution efficiency in lieu of separation efficiency (op. cit.). Moreover, some available polymers (e.g. alkyl-cellulose) have not been considered as yet in the data base and need to be included.

Significance to Biomedical Research and the Program of the ICD:

The program should be of the greatest usefulness for the practice of capillary zone electrophoresis in buffered polymer solutions, since it overcomes the arbitrariness of polymer choice and resulting inefficiency of separations.

Subproject (3) - Test and revision of the "Gaussian assumption" (Weiss, G. H., Sokoloff, H. and Chrambach, A.)

Objective:

The availability of band profiles from automated gel electrophoresis apparatus allows one to test the assumption of Gaussian band shape underlying the calculation of peak variance from band width, and of diffusion and dispersion coefficients. Objective of the study was to verify the accuracy of those two calculations and to revise them if necessary.

Major Findings:

The Gaussian shape of protein bands at the start of electrophoresis and at the lowest gel concentrations was verified. However, the band shape deviates progressively with increasing gel concentration and time.

Proposed Course:

The identical test must still be carried out for random-coiled macromolecules (SDS-proteins). The algorithms which assume Gaussian band shape will be revised according to the data, and mathematical models accounting for the deviation from Gaussian band shape will be sought and tested.

Significance to Biomedical Research and the Program of the ICD:

Accurate knowledge of the dispersion of macromolecules in gels and polymer solutions will allow for prediction of resolution, and thereby upgrade the efficiency of electrophoretic separations.

Subproject (4) - Prediction of the electroelution time required for near-quantitative recovery of proteins (Aldroubi, A., Zakharov, S. F. and Chrambach, A.)

Objective:

To exploit the preparative capacity of automated apparatus, it is desirable to know the electroelution time required for a macromolecule of known mobility. A computer program providing that time is the most user-friendly way for achieving that prediction.

Major Findings:

A program was created which interpolates a data base of SDS-protein recoveries at particular electroelution times. It predicts the necessary time for near-quantitative recovery of unknown SDS-proteins. The prediction for such recovery made by the program was verified experimentally to a first approximation, i.e. within 10-30%.

Proposed Course:

The data base of recoveries of SDS-proteins needs to be enlarged to improve the accuracy of the predicted electroelution time. Equivalent data bases for globular proteins and DNA are needed to allow for application of the program to those conformational types. The program input should be extended to include dispersion data.

Significance to Biomedical Research and the Program of the ICD:

Prior to automated gel electrophoresis apparatus, recovery in preparative electrophoresis was unmonitored and unpredictable. Remedy of that problem promises widespread applications of preparative automated apparatus in biochemistry.

Subproject (5) - Horizontal gel electrophoresis with large sample volume capacity (Chang, H.-T. and Chrambach, A.)

Objective:

Prior to this work, the sample load volume limit of horizontal gel electrophoresis had been 0.01 to 0.05 ml, requiring a preconcentration step with necessary losses in applications to the dilute solutions of macromolecules frequent in biology. The load limitation was particularly vexing in preparative applications. This study attempted to solve the problem, using suitable sample loading technique and sample stacking by application of discontinuous buffer systems.

Major Findings:

Sample application through a funnel shaped applicator of 17 ml volume containing 3 ml of stacking gel allows a 1.5 ml sample to be concentrated into a narrow starting zone prior to entrance into the resolving gel. Resolving gel patterns on horizontal gels of 3mm thickness derived from 0.5, 1.0 and 1.5 ml are indistinguishable.

Proposed Course:

To be widely applicable, the loading of large volume samples still needs to be demonstrated in discontinuous buffer systems spanning a wide range of operative pH's. Moreover, since perfect contact between the stacking gel in the funnel and the application zone in the resolving gel is necessary, a wide range of gel concentrations and gel types need to be evaluated as well as the question whether stacking and resolving gels need to be of the same gel type.

Significance to Biomedical Research and the Program of the ICD:

Large load volume capacity of horizontal gels will benefit the practice of gel electrophoresis by i) avoiding losses in preconcentration steps, and ii) augmenting by 1-2 orders of magnitude the preparative capacity of such gels.

Subproject (6) - Electrophoresis in polymer solutions, using automated gel electrophoresis apparatus (Chang, H.-T. and Chrambach, A.)

Objective:

Conventionally, electrophoresis of subcellular-sized particles (which are excluded from gels) in buffered polymer solutions uses capillary zone electrophoresis apparatus which lacks preparative capacity. The need for preparative electrophoresis of microsome components (see joint report) prompted us to investigate the applicability to polymer solutions of automated gel electrophoresis apparatus which does have preparative capacity.

Major Findings:

Subcellular-sized polystyrene particles form bands in electrophoresis on the automated gel apparatus using buffered polymer solutions of various types as the "molecular sieve". The degree of homogeneity of the peaks differs among polymer types and increases with polymer concentration. A syringe device for withdrawing of the peak for preparative purposes was designed and

shown to provide yields of 70-90% by sequential withdrawal from single peaks. A mature design is being constructed at this time.

Proposed Course:

The major remaining question is to what degree bands can be separated by electrophoresis in polymer solutions within the limits of field strength available from the Peltier-cooled automated apparatus with only 20 Watt heat dissipation capacity. To test that question, the technique will be applied to pairs of polystyrene particles differing to varying degrees in molecular weight. Once the capacity of the system for separations is known, it will be applied to the separation of microsome components, first analytically, then preparatively.

Significance to Biomedical Research and the Program of the ICD:

If capable of significant separations, the system will be widely applicable to the isolation of subcellular-sized particles with quantitative monitoring of recovery.

Subproject (7) - The reproducibility of mobilities in gel electrophoresis (Zakharov, S. F., Chang, H.-T. and Chrambach, A.)

Objective:

Although the entire edifice of "quantitative electrophoresis" is based on the measurement of the mobility of bands, no previous study has determined its precision, presumably in view of the labor intensiveness of large numbers of mobility measurements. Automated apparatus with an output of mobility at desired scanning intervals during electrophoresis in 8 parallel channels has changed that situation. Routinely, up to 30 mobility values per run per channel are automatically available and have been collected for statistical analysis.

Major Findings:

Absolute mobilities of globular or random-coiled proteins in agarose gel electrophoresis at all concentrations obtain with a precision of less than 3% standard deviation intraexperimentally, and slightly more than 3% interexperimentally. Native protein mobility in polyacrylamide gel varies within a standard deviation of 7% in one case, but in application to another protein rises progressively above that value at gel concentrations above 10 %T.

The precision of relative mobilities of DNA fragments does not exceed that of the absolute mobility stated above. Discontinuous buffer systems fail to give a constant mobility due to the progressive replacement of a less conductive leading buffer phase by a more conductive trailing buffer phase until the moving boundary has reached the end of the migration path. Mobility values are indistinguishable whether measured automatically or manually. A byproduct of the study of precision has been the observation that the fluorescent labeling of the particle required by the optical detection system of the automated apparatus induces in some proteins a mobility that exceeds that obtained from the SDS-derivatization of the fluorescently labeled protein, suggesting a competition for cationic sites on the protein surface between fluorescenating agent and SDS.

Proposed Course:

"Quantitative electrophoresis" based on the measurement of mobility at several gel (or polymer) concentrations provides information that is not available from the conventional visual interpretation of band patterns on gels of a single concentration. That information includes i) molecular size in terms of the radius (nm) of an equivalent sphere; ii) gross conformational features such as those due to molecular bending; iii) molecular malleability; iv) geometric mean radii of particles up to the micron-range; v) estimates of surface net charge; vi) a test of molecular identity based on a statistical evaluation of size and net charge; vii) gel fiber radius and length per unit volume.

For anyone contemplating to exploit that information regarding particle size and gel fiber properties that is available from "quantitative" gel electrophoretic methods based on mobility, the first question that need to be satisfied concerns the reproducibility of mobility. That question has been answered with respect to a few proteins and a few gel variants. A vastly larger data base is still required before the claim of sufficient reproducibility of mobility between experiments can be generalized. The proposed course of the project is, therefore, to collect and statistically evaluate the large number of mobility values that will accrue in the course of automated gel electrophoresis over the next few years. It will not be necessary to do any experiments with the sole purpose of measuring mobility.

Significance to Biomedical Research and the Program of the ICD:

The significance of the reproducibility of electrophoretic mobility in a particular milieu (pH, ionic strength, temperature) is that it

gives the green light to accessing the information available from "quantitative electrophoresis" as embodied in the PAGE-PACK and ELPHOFIT computer programs.

Subproject (8) - Advances in the mechanistic interpretation of "molecular sieving" (Radko, S. P. and Chrambach, A.)

Objective:

The long-term objective is an understanding in physico-chemical and polymer-chemical terms of the mechanism which give rise to separation on the basis of size and shape differences between macromolecules. Axiomatically, the efficiency of "size separation" methods should improve to the degree of that understanding.

Major Findings:

The previous report dealt with the observation that the relation between retardation in electrophoresis and particle size was triphasic when studied over a sufficiently wide range of sizes. The first 2 phases of that relationship were interpreted as being due to a collision mechanism between large or immobilized polymer chains and a relatively small particle in the 1st phase, followed by a displacement mechanism between flexible polymers and relatively large particles in the 2nd phase. The likely cause for the existence of the 3rd phase is the increased rigidity of very long polymer chains as measured by their increased relaxation times. Relaxation times were estimated on the basis of the mobility response to varying the field strength. Further new findings in that area are that i) retardation varies in triphasic fashion as a function of the molecular weight of the polymer as well as that of the particle; ii) band width also varies with polymer or particle size in the same fashion; iii) the monotonic decrease of entanglement threshold with polymer weight shows that it cannot be the decisive cause of retardation which relates to polymer weight in 3 phases; iv) retardation is maximal at a particular particle or polymer size, which shows that molecular sieving cannot be accounted for by the particle/polymer size ratio; v) mobility and band width change with concentration of a smaller polymer at different rates, with concentration of a larger polymer at the same rate.

Proposed Course:

We will attempt to incorporate the numerous measurements and observations reported above into a mathematical model of molecular sieving and to test the model experimentally. Since the dependence of retardation in polymer networks on particle size is triphasic, the data base for such a model must comprise specifications of parameters such as band width, entanglement threshold derived from viscosity data and field strength dependence of retardation to be able to predict the optimal polymer for separation within a given range of particle sizes.

Significance to Biomedical Research and the Program of the ICD:

A mathematical model of retardation in the electrophoresis using polymer solutions should allow us to design and predict capillary zone electrophoresis separations and thereby upgrade their effectiveness.

Publications:

Aldroubi A, Zakharov SF, Chrambach A. A computer program for predicting recovery of SDS-protein in the automated HPGE-1000 apparatus, Appl Theor Electrophoresis 1995, in press.

Buzas Z, Wheeler DL, Garner M M, Tietz D, Chrambach A. Transverse pore gradient gel electrophoresis, using the PhastSystem, Electrophoresis 1994;15:1028-1031.

Chang H-T, Chrambach A. Dispersion coefficients of a protein and DNA fragment in polyacrylamide gel electrophoresis as a function of parameters defining the effective gel pore size and particle size, Electrophoresis 1995;16:895-898.

Chang H-T, Chrambach A. Electrophoresis of a subcellular-sized particle in polymer solutions, using automated horizontal gel apparatus, Appl Theor Electrophoresis 1995, in press.

Chang H-T, Chrambach A. Horizontal gel electrophoresis with sample volumes up to 1.5 ml, using a discontinuous buffer system and automated apparatus, Electrophoresis 1995, in press.

Chrambach A, Wheeler DL. Capabilities and potentialities of transverse pore gradient gel electrophoresis, Electrophoresis 1994;15:1021-1027.

Chrumbach A, Yarmola E, Zakharov SF, Garner MM. Commercial automated gel electrophoresis apparatus: Application to DNA, band dispersion, nonlinear Ferguson curves, and isolation. Electrophoresis 1995;16:713-718.

Pulyaeva H, Zakharov SF, Garner MM, Chrumbach A. Detection of a single base mismatch in double-stranded DNA by electrophoresis on uncrosslinked polyacrylamide gel, Electrophoresis 1994;15:1095-1100.

Radko SP, Chrumbach A. Molecular sieving by polymer solutions: Dependence on particle and polymer size and dispersion, independence of polymer entanglement, Appl Theor Electrophoresis 1995, in press.

Radko SP, Garner MM, Caiafa G, Chrumbach A. Molecular sieving of polystyrene carboxylate of a diameter up to 10 um in solutions of uncrosslinked polyacrylamide of Mr 5×10^6 using capillary zone electrophoresis, Anal Biochem 1994;223:82-87.

Weiss GH, Garner MM, Yarmola E, Bocek P, Chrumbach A. A comparison of the resolution of DNA fragments between agarose gel and capillary zone electrophoresis in agarose solutions, Electrophoresis 1995, in press.

Wheeler DL, Chrumbach A. A computer simulation accounting for dissimilar electrophoretic behavior between two similarly curved DNA fragments due to a difference in arc-length, Electrophoresis 1994;15:885-889.

Wheeler DL, Chrumbach A. Computer simulation of the directional displacement of rod-shaped, arc-shaped and circular objects in an array of obstacles, representing a simple model for the gel electrophoresis of small DNA, Biopolymers 1995;35:179-185.

Yarmola E, Chrumbach A. Band width measurement in automated gel electrophoresis apparatus: DNA dispersion in a discontinuous system and in a single buffer, Electrophoresis 1995;16:345-349.

Yarmola E, Chrumbach A. Nonlinear "Ferguson curves" by two runs of the commercial automated HPGE-1000 gel electrophoresis apparatus with intermittent scanning of fluorescence, Electrophoresis 1995;16:350-353.

Zakharov SF, Chrambach A. The relative separation efficiencies of highly concentrated, low crosslinked polyacrylamide gels compared to conventional gels of moderate concentration and crosslinking, Electrophoresis 1994;15:1101-1103.

Zakharov SF, Garner MM, Chrambach A. Fluorescent labeling of DNA with ethidium homodimer without measurable decrease in DNA mobility: Application to automated apparatus, Anal Biochem 1995;224:195-198.

Zakharov SF, Garner MM, Chrambach A. Recovery of SDS-protein and DNA using commercial automated gel electrophoresis apparatus, Appl Theor Electrophoresis 1995, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01400-13 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Dynamics of the Growth and Development of Bone

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: A. Yergey Head SMAMS, LTPB, NICHD

Others: N. Vieira Biologist SMAMS, LTPB, NICHD

COOPERATING UNITS (if any)

Human Genetics Branch, NICHD (J. Marini); Child Nutrition Research Center, Houston, TX (S. Adams); Nemours Children's Clinic, Jacksonville, FL (N. Mauras)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Metabolic Analysis and Mass Spectrometry

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☒ (a) Human subjects ☐ (b) Human tissues ☐ (c) Neither
☒ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

1. A pilot study in a surgically post-menopausal female was carried out using the naturally occurring calcium radio-isotope, ^{41}Ca (half life 10^5 yr). Accelerator mass spectrometric measurements, made in collaboration with Stewart Freeman of Lawrence Livermore National Laboratory, of ^{41}Ca were coupled with results from a typical stable isotope dilution study carried out simultaneously. Values of calcium turnover were calculated from both the stable and ^{41}Ca tracer studies were calculated to be 220 and 234 mg/d respectively. These results are a successful demonstration of the ability to obtain comparable results from steady-state and dynamic approaches in a complex protocol. This validation will increase confidence in the use of the steady state approach for studies of calcium skeletal dynamics at long times after the administration of the ^{41}Ca tracer, i.e., the subject's lifetime and has led directly to planning of more medically significant studies. 2. The effects of estrogen replacement on calcium metabolism in 6 hypogonadal females, 5 with Turner's Syndrome and one with hypogonadotropic hypogonadism in paired pre- and post-treatment studies of Ca metabolism. We found a significant increase in whole body Ca retention, V_{bal} changing from 3.6 ± 1.3 to 10.2 ± 9.2 mg/d, $p=0.04$, and a significant decrease in whole body turnover from 74.6 ± 11.7 to 65.2 ± 9.7 mg/d, significant, $p=0.04$. Based on the results of this work, we speculate that estrogen, even at the virtually unmeasurable levels at which it is secreted from the immature ovary of normal pre-pubescent females, is extremely important for the development of a normal skeleton. The absence of these low levels of hormone during early childhood is closely associated with retardation in normal skeletal development. 3. We devised a method for estimating the fraction of urinary calcium which arises from recent dietary sources. In a study of 38 healthy female children, we have concluded that the contribution of calcium from recent dietary sources is trivial and about averages 8% of absorbed Ca.

Project Description

Objectives:

The principal objective of this work is to determine the dynamics by which calcium is accreted to and removed from bone during human growth and development. The investigations conducted in this project have represented an attempt to assemble congruent, or at least convergent, views of calcium dynamics from studies of intact human subjects. To date this objective has been met by studies of absorption and distribution of calcium in healthy infants and children, in children with diseases suspected of altering calcium metabolism and in healthy women in altered states of calcium metabolism such as pregnancy and lactation.

Methods Employed:

The principal methodology used employs stable isotopes of calcium as tracers for the entrance and distribution of calcium in the body. Determination of isotope balances allows the calculation of important parameters of calcium metabolism, principally the rates of absorption and excretion. Fitting of mathematical models to the time-dependence of the disappearance of the intravenous isotope, when coupled with the balance data, leads to the calculation of the rates of calcium accretion to and resorption from bone. Thermal ionization mass spectrometric measurements of calcium isotope ratios of material isolated from blood and excreta are the observations obtained from these studies. The mathematical models which are fit to these observations represent ongoing attempts to characterize the dynamics of bone remodeling.

Major Findings:

1. ^{41}Ca and Accelerator Mass Spectrometry. There are number of basic issues that bear on the body of studies that have been performed using isotope dilution approaches. These issues can be reduced to a question of the size or fraction of the skeletal mass that is involved with the exchange of tracer over the 1-2 week period typically involved in these studies. We have performed a pilot study using the naturally occurring calcium radio-isotope, ^{41}Ca (half life 10^5 yr), in which a lifetime dose was administered to a human subject in a manner that indicates to date that there will be a lifetime measurable level of the isotope in the subject's skeleton. Accelerator mass spectrometric measurements, made in

collaboration with Stewart Freeman of Lawrence Livermore National Laboratory, of ^{41}Ca were coupled with results from a typical stable isotope dilution study carried out simultaneously. Values of calcium turnover were calculated from both the stable and ^{41}Ca tracer studies were calculated to be 220 and 234 mg/d respectively. Since the basis of the ^{41}Ca measurements is a series of steady state balance determinations that depend on a knowledge of absorption, while the stable isotope measurements use a dynamic model of internal distribution, this study has validated the use of the steady-state approach. This will lead to confidence in the results of studies of calcium skeletal dynamics at long times, i.e., the subject's lifetime, after the administration of the ^{41}Ca tracer. This successful demonstration of comparable kinetics is also a successful demonstration of the execution of a complex protocol involving complex sample preparation for the ^{41}Ca samples. This dual success has led to the planning of an extensive study of long term calcium kinetics in peri- and post-menopausal women.

2. In collaboration with Dr. Nelly Mauras, Division of Endocrinology, Nemours Clinic, Jacksonville, FL, we have investigated the effects of estrogen replacement on calcium metabolism in 6 hypogonadal females, 5 with Turner's Syndrome and one with hypogonadotropic hypogonadism. Following a baseline Ca dynamics study, each subject was treated with ethinyl estradiol for 4 weeks, and after the passage of an additional 4 weeks, a second Ca dynamics study was performed. We found that, in addition to a significant increase in whole body Ca retention, V_{bal} changed from 3.6 ± 1.3 to 10.2 ± 9.2 mg/d, significant at $p=0.04$, whole body turnover decreased significantly from 74.6 ± 11.7 to 65.2 ± 9.7 mg/d, significant at $p=0.04$. We conclude that the estrogen treatment results in a change in Ca metabolism leading to what may be a recouping of mineralization losses normally seen in these subjects and that consideration of earlier treatment with small estrogen doses designed to prevent low bone mineral density should be considered for these patients. A more important idea to be derived from this work is a consequence of the fact that the subjects of this study are congenitally absent of any estrogen sources. Based on the results of this work, we speculate that estrogen, even at the virtually unmeasurable levels at which it is secreted from the immature ovary of normal pre-pubescent females, is extremely important for the development of a normal skeleton. The absence of these low levels of hormone during early childhood is closely associated with retardation in normal skeletal development.

3. In a collaborative study with Dr. Thomas R Welch, Dept. of Pediatrics, Division of Nephrology, Children's Hospital, Cincinnati, OH, we devised a method for estimating the fraction of urinary calcium which arises from recent dietary sources. From a study of 38 healthy female children, we have concluded that the contribution of calcium from recent dietary sources is trivial and averages about 8% of absorbed calcium. This observation is entirely consistent with our best current model of calcium distribution throughout the body in which it is envisioned that the short-term depot of calcium is labile stores in bone. The implications of this work have potentially important consequences for current treatments of urolithiasis, insofar that restriction of dietary calcium may be an inappropriate intervention.

Significance to Biomedical Research and the Program of the ICD:

An overall understanding of the developmental aspects of calcium absorption, excretion and distribution to bone as a function of development in humans is emerging as a consequence of these studies. Our results can begin to be used to interpret abnormalities of calcium metabolism in a manner that can be therapeutically useful.

Proposed Course:

The nature of this project has become progressively more clinical in what is fundamentally a physical chemistry oriented section. As a consequence, the section will move away from this area of research by completing current projects, by not undertaking new collaborative efforts and, to the extent possible, by finding physical-measurement-oriented clinical colleagues interested in continuing current lines of investigation.

Clinical Protocols:

Project Title: Kinetics of Calcium Metabolism
Project No.: 91-CH-137

Project Title: Introduction of a Lifetime Measureable ^{41}Ca Dose Into the Body of a Single Subject and Direct Measurement of V_o and V_{o-} .
Project No.:

Publications:

Abrams SA, Schanler RJ, Yergey AL, Vieira NE, Bronner F. Compartmental analysis of calcium metabolism in very low birth weight infants, *Pediatr Res* 1994;36:424-428.

Abrams SA, Vieira NE, Yergey AL. Interpretation of stable isotope studies of calcium absorption and kinetics. In: Subramanian S, Wastney ME, eds. *Mathematical Models of Trace Metal Nutrition*. Boca Raton: CRC Press, 1995;283-290.

Abrams SA, Yergey AL, Heaney RP. Relationship between balance and dual tracer isotopic measurements of calcium absorption and excretion, *J Clin Endocrinol Metab* 1994;4:965-969.

Ebeling PR, Yergey AL, Vieira NE, Burritt MF, O'Fallon WM, Kumar R, Riggs, BL. Influence of age on effects of endogenous 1,25-dihydroxyvitamin D on calcium absorption in normal women, *Calcif Tissue Int* 1994;55:330-334.

Goans RE, Abrams SA, Vieira NE, Marini JC, Perez MD, Yergey AL. A three hour measurement to evaluate bone calcium turnover, 1995;16:33-38.

Goans RE, Abrams SA, Vieira NE, Yergey AL. Frequency response of the calcium control system. In: Subramanian S, Wastney ME, eds. *Mathematical Models of Trace Metal Nutrition*. Boca Raton: CRC Press, 1995;333-340.

Goans RE, Weiss GH, Abrams SA, Perez MD, Yergey AL. Calcium turnover kinetics show decreased irreversible flow to bone in glucocorticoid treated patients, *Calcif Tissue Int* 1994;55:330-334.

O'Brien KO, Allen LH, Quatromoni P, Siu-Caldera M-L, Vieira NE, Perez A, Turner A, Holick MF, Yergey AL. High fiber diets slow bone turnover in young men but have no effect on efficiency of intestinal calcium absorption, *J Nutr* 1993;123:2122-2128.

Vieira NE, Yergey AL. Extraction of serum and urine calcium using ion exchange membrane filters for isotope enrichment analysis using thermal ionization mass spectrometry, *Anal Chem* 1995, in press.

Weiss GH, Goans RE, Gitterman M, Abrams SA, Vieira NE, Yergey, AL. A non-markopvian model for calcium kinetics in the body, *J Pharmacokinet Biop* 1994;22:367-379.

Welch TR, Abrams SA, Shoemaker L, Yergey AL, Vieira NE. Precise determination of the absorptive component of urinary calcium excretion using stable isotopes, *Pediatr Nephrol* 1995;9:295-297.

Yergey AL. Issues in constant tracer infusion and mineral metabolism. In: Allen LH, King JC, eds. *International Conference on Nutrient Regulation During Pregnancy, Lactation and Infant Growth*. New York: Academic Press, 1995;279-290.

Yergey AL, Abrams SA, Vieira NE, Esteban NV, Bronner F. Calcium metabolism in pregnancy and lactation - An individual study. In: Subramanian S, Wastney ME, eds. *Mathematical Models of Trace Metal Nutrition*. Boca Raton: CRC Press, 1995;29-36.

Yergey AL, Vieira NE, Abrams SA, Marini J, Goans RE. Use of stable isotopic tracers in studies of whole body calcium metabolism, *Connective Tissue Res* 1995, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01408-04 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Stability and Specificity of DNA-Protein Interactions

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	M. Garner	Expert	LTPB, NICHD
Others:	A. Chrambach	Head	SMA, LTPB, NICHD
	J. Zimmerberg	Chief	LTPB, NICHD

COOPERATING UNITS (if any)

Laboratory of Biochemistry and Metabolism, NIDDK (D.C. Rau)
Laboratory of Molecular Biology, NCI (S. Adhya)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Macromolecular Analysis

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

This project has been terminated.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01409-10 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Membrane Transport and Fusion

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	J. Zimmerberg	Chief	LTPB, NICHD
Others:	I. Plonsky	Visiting Fellow	LTPB, NICHD
	E. Suss-Toby	Visiting Fellow	LTPB, NICHD
	P. Bronk	IRTA Fellow	LTPB, NICHD
	V. Ratnov	Adjunct Scientist	LTPB, NICHD
	E. Ratnov	IRTA Fellow	LTPB, NICHD
	M. Cho	Biologist	LTPB, NICHD
	M. Tahara	Visiting Fellow	LTPB, NICHD
	G. Ward	Senior Staff Fellow	LMR, NIAID

COOPERATING UNITS (if any)

Laboratory of Malaria Research, NIAID

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Membrane and Cellular Biophysics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

4.5

PROFESSIONAL:

4.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Membrane topology defines life: the inside of a cell is highly organized at the molecular level, at almost solid-state densities; the outside environment of cells is usually more chaotic. To understand the mechanisms by which cells control their membranes' topologies, we have continued to develop new methods and use new and old techniques to study two aspects of cell membrane remodeling. A new algorithm for cell capacitance measurement has been developed. This algorithm uses calibration jumps caused by a known change in the compensation circuit parameters of the patch clamp amplifier to determine the phase shift and artifactual attenuation, allowing automatic phase angle adjustments and reconstruction of the undistorted admittance values of circuits formed by fusing cells and calculation of the electrical parameters of a single cell. We used this software to measure the electrical properties of fusion induced by two similar but distinct proteins, GP64 for the Op and Ac strains of baculovirus. The conductance of the Op-mediated initial fusion pore varies within a range 0.3 - 1.9 nS with a mean of 0.7 ± 0.4 nS. This differs significantly from the Ac-induced pore conductance: 1.0 ± 0.3 nS. This lays the foundation for structure/function studies of fusion pore phenotype. In the second membrane event, endocytosis, we are exploring the electrical manifestations of parasite invasion, which culminates in the pinching-off of a parasitic vacuole (PVM). The application of simultaneous optical and capacitance measurements during parasite invasion enabled us for the first time to show that there is no significant change in host cell surface area during invasion of *Toxoplasma gondii*. Thus the parasite does not contribute a significant amount of lipid to the PVM formation. In addition, pinching off of the PVM from the host cell membrane is seen in each case as a sudden 0.2-0.3 pF drop in C_m . A fission pore forms during the pinching off process of the PVM. The parasitic fission pore is the endocytotic equivalent to the fusion pore. The parasitic fission pore exhibits flicker and semistable stages with an average conductance of 2 nS. The pore reaches a conductance of less than 1 nS just before closing. Correlated optical and electrical records of more than 100 invasions have shown that transient changes in membrane current (spike) occur at precisely the moment that the host cell is first contacted by the apical end of the parasite.

Project Description:Objectives:

Membrane topology defines life: the inside of a cell is highly organized at the molecular level, at almost solid-state densities; the outside environment of cells is usually more chaotic. To understand the mechanisms by which cells control their membranes' topologies, we have continued to develop new methods and use new and old techniques to study two aspects of cell membrane remodeling. In the first, membrane fusion, we have focused on the fusion pore which initiates the aqueous continuity that marks fusion. In the second, endocytosis, we are exploring the electrical manifestations of parasite invasion, which culminates in the pinching-off of a parasitic vacuole.

For methods, our goals were to develop new and to enhance existing algorithms for the calculation of different electrical parameters of fusing cells. We also aimed at devising a computer simulation of the fusion process, implementing of these algorithms into computer programs, improving the performance and user interface of existing software for membrane current acquisition and cell admittance calculation, visualization and analysis, and integrating programs developed for different computer interfaces and different research purposes into a universal MS Windows based software package.

For fusion pores, our goal is the elucidation of the mechanisms by which fusion proteins cause fusion. Viral fusion proteins are unique in being available for direct study at this time. GP64, a major envelope protein, was identified as the fusion protein of baculovirus (Blissard and Wenz, 1992, J. Virol., 66:6829-35). Our strategy for determining how a fusion protein works is to compare fusion pore phenotype with fusion protein sequence. We will begin with a comparison of two similar fusion proteins, the GP64 of two different baculoviruses. We then can devise chimera of the two proteins to determine which domains contribute to fusion pore phenotype. Fusion proteins contain hydrophobic domains or "fusion peptides" (FP), which interact with host cell membrane during fusion. FP are typically conserved within but not between virus families. Amino acids 223-228 of *Orgyia pseudotsugata* multiple nuclear polyhedrosis virus (Op) GP64 was characterized recently as the fusion peptide (Monsma and Blissard, 1995, J. Virol., 69:2583-95). The FP of another virus from *Baculoviridae*, *Autographa californica* (Ac), has the same sequence except Ala223 replaces 223Val (Whitford et al., 1989, J. Virol., 63:1393-99).

To understand membrane infolding, we are characterizing for the first time conductance and capacitance changes which accompany invasion of the parasite *Toxoplasma gondii* into COS cells. This is a powerful new approach for studying membrane events associated with host-parasite interaction. These experiments may shed light on the mechanism by which the parasite invades and becomes enveloped by a membrane known as the parasitophorous vacuole membrane (PVM). The PVM pinches off from the host cell membrane at the completion of invasion. Therefore, invasion of host cells by parasites represents a novel model system for studying the biochemistry and electrophysiology of endocytosis and endocytotic vesicle fission.

Method Employed:

Software development: The VISUAL C++ 1.52 development system and the Microsoft Windows 3.1 Software Development Kit were the compilers and interface tools used to create these programs. For interaction with Microsoft Access during the maintenance of the database of experimental parameters we used ODBC (Open Data Base Connectivity) libraries and API (application programming interface), version 2.0. For communication between different program components of the software package we used DDE (dynamic data exchange) protocol and DDEML (dynamic data exchange management library).

We used object-oriented design and methodology in order to reduce the time of software development, to simplify further modifications and to ensure the software portability. Some of the developed program components (hardware related routines and most frequently used dialog boxes) were put into DLL's (dynamic link libraries) to enable their sharing between applications incorporated into the package. Slopes of segments of data were analyzed by regression analysis. Our approach to noise processing was based on a calculation of standard deviations and usage of percentiles for the Gaussian distribution.

Viral fusion electrophysiology: Ld insect cells were infected with Op baculovirus at 3-5 MOI (units of multiplicity of infection) and studied 36-48 hours later. Fusion pore conductance (G_p) was calculated using time-resolved admittance measurement techniques. A 1 kHz, 50 mV peak-to-peak sine wave was superimposed on the holding potential (-30 mV). Prior compensation of R_s (resistor in series with cell circuit) and C_{m1} (first cell capacitance) and adjustments of the phase angle were done as described earlier (Joshi and Fernandez, 1988, Biophys. J., 53: 885-892). Sinusoidal current was filtered at 5 KHz, acquired at 40 KHz and separated "on-line" into imaginary ($\text{Im}(Y)$), real ($\text{Re}(Y)$) and direct current (G_{DC}) components of pipette-cells admittance (Y) by the computer program described above. G_p was calculated on the basis of the equivalent circuit (from Zimmerberg et al., 1987, PNAS, 84:1585-1589) as $G_p = (\Delta \text{Re}(Y)^2 + \Delta \text{Im}(Y)^2) / \Delta \text{Re}(Y)$, where $\Delta \text{Re}(Y)$ is the increment of the real [$\text{Re}(Y) = G_{m1} + (\omega \times C_{m2})^2 / G_p \times (1 + (\omega \times C_{m2} / G_p)^2)$] and $\Delta \text{Im}(Y)$ is the increment of the imaginary [$\text{Im}(Y) = \omega \times C_{m2} / (1 + (\omega \times C_{m2} / G_p)^2)$] component of Y ; G_{m1} is the conductance of a cell which is connected, C_{m2} is the capacitance of a cell which is not connected to the patch pipette, $\omega = 2\pi f$, and f is the sinewave frequency. To trigger cell-cell fusion, acid solution was ejected from a micropipette, placed at about one cell diameter aside a cell pair. Ejections were performed using a PV830 Pneumatic Picopump (WPI, Sarasota, Florida). A valve-operating voltage pulse of the picopump was also acquired and stored together with $\text{Re}(Y)$, $\text{Im}(Y)$ and G_{DC} data points. Kinetics of fusion was studied by measuring the delay between trigger and fusion pore appearance. AcMN PV-induced fusion in SF9 cells was characterized earlier.

Cell cultures for parasitology: COS 1 cells (African green monkey SV40 transformed, fibroblast-like kidney cell line) were obtained from the ATCC and were cultured in DMEM complete at 37°C. Trypsinized cells were used (2 min. in Trypsin-EDTA). *Toxoplasma gondii* RH

strain tachyzoites were cultured in monolayer of human foreskin fibroblasts (ATCC #HS68). Before use, parasites were passed twice through a 27 gage needle, filtered through 10 μ m PCTE filters (Poretics), spun for 5 min. at 1100 x g and resuspended in the bath solution. Experiments were performed at 35°C, using a Bioptics heating system.

Parasite invasion electrophysiology: Sine waves of 80 mV amplitude and 200 Hz and/or 1000 Hz frequencies were given by an Axolab-1 computer interface (Axon Instruments Inc.). An EPC7 (List Electronics) amplifier was used for converting to voltage the output current of whole-cell clamped cells. Data was filtered by a 3 kHz filter, and digitized at 31.25 μ s rate, and processed with the software documented above. The bath solution containing: 118 mM NaCl, 5.85 mM KCl, 1.2 mM MgCl₂, 2.52 mM CaCl₂, 5mM Hepes, and 11 mM glucose, titrated to pH=7.4 with NaOH. Recordings were performed either under whole cell configuration or perforated patch. The pipette solution for whole-cell recordings was based on KCl or KAspartate and contained: 122 mM KCl or KAspartate, 2 mM MgCl₂, 11 mM EGTA, 1 mM CaCl₂, and 5 mM Hepes, titrated to pH=7.26 with KOH. 1 mM ATP and 1 mM GTP were added just before use. The pipette solution for the perforated patch contained saturated amphotericin B. 1.5-2 MOhm resistance pipettes were used. The pipettes were made from WPI thin wall glass with filament (TW150F-4) on a Narishige vertical puller (model PB-7).

Optical measurements for parasite invasion: Nomarski and fluorescent images were collected by a Dage camera (MTI-VE1000SIT), stored on a Sony U-matic videocassette recorder (V0-5800) and processed by a Videoscope image processor. A RCA screen splitter/insertor was used to combine electrical and optical information on the same videoframe. Fluorescein-conjugated antibody against SAG1, a major parasite surface protein (Biogenex) was used to assay invaded parasites versus external ones.

Major Findings:

Capacitance Software: A new algorithm for cell capacitance measurement has been developed. This algorithm uses calibration jumps caused by a known change in the compensation circuit parameters of the patch clamp amplifier to determine the phase shift and attenuation introduced by experimental equipment. This approach allows one to **reconstruct the undistorted admittance values of a circuit formed by fusing cells and to calculate the electrical parameters of a single cell**, using multi-frequency sine wave stimulation. The existing algorithms for pore calculations have been modified to work with the reconstructed circuit admittance.

New software for multi frequency lock-in calculations based on burst mode acquisition has been developed. This software uses an original approach for the synchronization of stimulus and the cell response based on the TTL acquisition triggering and allows simultaneous data acquisition and processing. This new program has advantages and new features compared with the old software used in the laboratory:

- 1) it has a convenient MS Windows graphical user interface.
- 2) it allows one to customize the length of the acquisition buffer which in turn improves the accuracy of event markers and the general sensitivity of the program.
- 3) it eliminates previous limitations on the stimulating frequencies and the number of acquired points per sine wave period, thus allowing an **increase in the signal to noise ratio**.
- 4) it can run in the background, allowing simultaneous processing and browsing of other sets of experimental data, as well as entering compensation circuit parameters on the fly.
- 5) while running in the background, the program can perform **automatic phase angle adjustments** for all stimulating frequencies.

The software has been incorporated into an integrated MS Windows-based package for electrical data processing and analysis, which is being developed in the laboratory. This analysis program (*Browse*) has the following new features:

- 1) maintenance of a database, keeping the main parameters of all experiments in a Microsoft Access DBMS (Data Base Management System) file. The database system format can be easily adjusted to the needs of different experimental groups. Currently the software supports formats for toxoplasma invasion and cell bilayer fusion experiments.
- 2) a broad range of changeable settings: various printing options (including the possibility of color printing), calibration parameters and color scheme customization.
- 3) an extended set of routines for experimental data conversion, including phase shifting, removal of irregular points, additional procedures for importing data, etc.
- 4) a broad range of new browsing routines, including vertical scrolling, "watch mode" allowing easy retrieving of channel and time values and their increments, a set of functions for working with bookmarks, etc.
- 5) an extended and enhanced set of procedures for calculation of different electrical parameters of fusing cells and fusion pores, based on the new algorithm outlined above, new subsets of admittance channels, and auxiliary procedures for direct calculation of admittance.
- 6) noise analysis procedures based on the outstanding points, calculation of means and standard deviations inside noise bounds.
- 7) calibration of signal and area calculations.
- 8) fitting data with slopes, estimation of the fitting quality, calculation of slope characteristics

and copying them to the clipboard for the successive usage in other scientific applications (Sigmaplot for Windows, etc.)

- 9) improved user interface, providing visual feedback on various user actions (selecting data intervals, mouse based scaling, etc.), simplified data browsing based on scrollbar usage, display of relevant electrical circuits and other new features.

A special installation program has been added to the software package. It allows quick and easy installation of the software and provides necessary version control of program components. A set of user documentation, including the description of all calculation routines has also been prepared.

Viral Fusion: The conductance (G_p) of the Op-mediated initial fusion pore varies within a range 0.3 - 1.9 nS with a mean of 0.7 ± 0.4 nS (Kolmogorov-Smirnov normality test: $P=0.4533$). **The Op-induced G_p differs significantly from the Ac-induced pore conductance** (1.0 ± 0.3 nS; two-sample Kolmogorov-Smirnov test was used for statistical evaluation). In experiments with Ac, the delay time to fusion after acidification (t_0) distribution was fit to the sum of two exponentials, yielding rate constants $k_1=0.68$ s⁻¹, $k_2=3.5$ s⁻¹. The t_0 distribution of Op was best fit to a single-exponential with a rate constant equal to 6.3 s⁻¹. Ac-induced fusion can be modeled with a three-state, two-transitional model ($A \rightarrow B \rightarrow F \rightarrow$) with $B_0/(A_0+B_0)=0.84$ (A_0 , B_0 are the concentrations of A and B elements at the moment of triggering). A two-state, one-transition kinetic scheme is adequate for Op-induced fusion.

Parasite Invasion: The application of simultaneous optical and capacitance measurements during parasite invasion enabled us for the first time to show that **there is no significant change in host cell surface area during invasion of *Toxoplasma gondii***. Cell Capacitance (C_m) stays relatively constant during the period of PVM formation. Statistical analysis on the C_m measurements during PVM formation suggests, in contrast to the predominant theory, the parasite does not contribute a significant amount of lipid to the PVM formation (if at all, between 0 to 10%). In addition, **pinching off of the PVM from the host cell membrane is seen in each case as a sudden 0.2-0.3 pF drop in C_m** . This result is consistent with the predicted size of the PVM ($40 \mu m^2$). We have often observed, during drops in capacitance, the "flicker" phenomena, which is reminiscent of the flickering increase in C_m seen during exocytosis. We have characterized the fission pore which forms during the pinching off process of the PVM. The parasitic fission pore is the endocytotic equivalent to the fusion pore (which forms during any exocytotic event). The parasitic fission pore exhibits semistable stages with an average conductance of 2 nS. The pore reaches a conductance of less than 1 nS just before closing. Unexpectedly, **there is a variable delay between the completion of constriction and the pinch-off event**. This delay, which lasts on average 139s ($n=16$), gives us a unique opportunity to study host and/or parasite factors which are necessary for membrane fission. Correlated optical and electrical records of more than 100 invasions have shown that **transient changes in membrane current (spike) occur at precisely the moment that the host cell is first contacted by the apical end of the parasite**. Our correlated measurements show that every invasion ($n>100$) is preceded by a spike but not

every spike is followed by invasion. Only actively moving parasites induced spikes. The number of spikes is correlated with number of active parasites delivered to the cell. Preliminary characterization of the spikes was done by measuring current under different DC holding potentials; the resultant I/V curve is linear, with a reversal potential of approximately -6 mV.

Significance to Biomedical Research and the Program of the ICD:

New methods for measuring basic cellular properties such as this new method for automating the measurements of cell surface area and permeability may have significant impact upon the measurements of cell properties in disease and pre-disease states. In our particular uses of this technology, two pathogenic states were studied. Enveloped viral infection is the source of much morbidity and mortality, from HIV, *influenza*, *herpes*, etc. *Toxoplasma gondii* is an important pathogen in pregnant women and immunocompromised persons, such as those infected with HIV. An understanding of the basic mechanisms underlying invasion by *Toxoplasma gondii* may lead to new chemotherapeutic or immune approaches to controlling the disease it causes.

Proposed course:

- 1) Further automate and improve our computer programs.
- 2) The finding of two strikingly different phenotypes of fusion pores produced by similar proteins (GP64) suggests a strategy for structure/function studies of the mechanisms of membrane fusion. While it is possible that the single amino acid difference in the fusion peptide region creates this modulation of membrane fusion, other factors, such as variations in the surface density of GP64, mutations outside the FP region, or differences between the membranes of the two cells may influence fusion pore phenotype. The measurements of fusion protein densities and experiments with cross-cellular expression of different mutant forms are essential for further analysis of the mechanism of membrane fusion.
- 3) Characterize the molecular mechanism underlying current transients (spikes): how are they induced, are they induced by a conductive element inserted by the parasite or does the parasite induce activation of a receptor-mediated channel? Another less likely option is that the parasite fuses temporarily with the host cell membrane before invading. It will be interesting to verify whether the spikes are a necessity for the invasion process. We have mutant parasites which lack specific proteins that are candidate for inducing spikes and will help determine the above questions.
- 4) Learn more about the lag preceding PVM pinch off. Can we use the delay to identify host or parasite factors involved in PVM pinch off? Is the protein, dynamin also involved in a non-clathrin dependent pinching-off process?

5) Characterize the fission pore in terms of life time, maximal conductance, minimum radius it reaches during flickering and before pinching off. We will compare it to the known features of the fusion pore. This study may shed light on the mechanisms that control vesicle fission and fission pore formation.

Publications:

Zimmerberg J, Blumenthal R, Curran M, Sarkar D, Morris S. Restricted movement of lipid and aqueous dyes through pores formed by influenza hemagglutinin during cell fusion, *J Cell Biol* 1994;127:1885-1894.

Melikyan GB, Niles WD, Ratinov VA, Karhanek M, Zimmerberg J, Cohen FS. Comparison of transient and successful fusion pores connecting influenza hemagglutinin expressing cells to planar membranes, *J Gen Physiol*, 1995, in press.

Zimmerberg J, Chernomordik L, Vogel S, Whalley T, Sokoloff A, Plonsky I, Chanturiya A. Intermediates in membrane fusion. *Cold Spring Harbor Symposia on Quantitative Biology*, 1995, in press.

10

11

12

13

14

15

16

17

18

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01415-05 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Kinetics of Exocytosis

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	J. Zimmerberg	Chief	LTPB, NICHD
Others:	P.S. Blank	IRTA Fellow	LTPB, NICHD
	M.S. Cho	Biologist	LTPB, NICHD
	S.S. Vogel	Senior Staff Fellow	LTPB, NICHD

COOPERATING UNITS (if any)

None

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Membrane and Cellular Biophysics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

We have continued our research on membrane fusion, the fundamental step in secretion, viral infection, fertilization and neuro-transmission. Sea urchin cortical granule exocytosis, an example of calcium triggered membrane fusion, has served as a model system. The sea urchin planar isolated cortex is the only model system that exists where the kinetics of exocytosis can be studied in the absence of reserve granule mobilization to docking sites or subsequent endocytotic activity. We have developed a two parameter model which describes both the kinetic and steady-state features observed in sea urchin cortical granule exocytosis. The kinetics of the fusion process are fit by our kinetic model and the maximal rate and extent of fusion reached a plateau at different calcium concentrations, in agreement with the model. We have determined from the calcium dependence of cortical granule exocytosis, and from the exposure time and concentration dependence of NEM inhibition, that 1) sea urchin egg cortical granule-plasma membrane docking sites have on average nine fusion complexes, 2) fusion complexes are randomly distributed among the population of exocytotic granules, and 3) calcium-secretion coupling is an intracellular manifestation of 'receptor reserve'. The relationship between docking/activation and subsequent calcium triggered fusion was examined by altering the pH of solutions matched for free calcium. Even though granules easily detach from the plasma membrane in the absence of calcium at pH greater than 7.5, no changes in the calcium activation curves were detected at pH 7.4.

Project Description:

Objectives:

The goal of this project is to elucidate molecular mechanisms involved in secretion, viral infection, fertilization, and neuro-transmission by studying triggered exocytosis. The merging of two membranes, membrane fusion, is a fundamental step in each of these biological processes. Sea urchin cortical granule exocytosis is an example of calcium triggered membrane fusion. The biophysical and biochemical properties of membrane fusion have been investigated in a model system composed of sea urchin cortical granules and egg plasma membrane. Steady-state and kinetic data have been analyzed using two model parameters, the average number and efficacy of activated fusion complexes.

The specific objectives of this project have been to 1) determine the number and distribution of fusion complexes at sea urchin egg cortical granule docking sites, 2) test a general, kinetic model which relates the rate and extent of triggered exocytosis to the number, distribution and efficacy of activated fusion complexes, and 3) determine if relationships between docking, activation and subsequent fusion are reflected in a model dependent interpretation of the rate and extent of fusion following alterations in physical parameters of the system.

Methods Employed:

Sea urchin egg cell surface complex, CSC, was prepared and assayed for calcium-triggered exocytotic activity according to established protocols. Fusion complexes were fractionally inactivated by mixing cell surface complex with solutions containing various concentrations of NEM (up to 20 mM) on ice. After the appropriate incubation time (up to 2 h), unreacted NEM was neutralized by the addition of 20 mM DTT. The calcium dependence of cortical granule fusion was determined using a turbidimetric microtiter-dish assay and solutions containing defined calcium concentrations. Survival curves were measured by monitoring light scattering from planar isolated sea urchin egg cortices that had been perfused with solutions containing defined concentrations of free calcium. A new chamber was designed, constructed and tested. Unidirectional flow of solutions resulted in the elimination of incomplete removal of cellular debris and the near elimination of bubbles. Vacuum filtering of solutions through a 0.22 μ m filter and low pass filtering have further improved the signal/noise characteristics of the light scattering assay. Changes in the light scattering signal are now resolvable during and immediately following rapid solution changes on the 10's of msec time scale. The fraction of refractile granules was measured using planar isolated sea urchin egg cortices. Granules were imaged using a Plan-Neofluar 63X 1.25 NA objective and Nomarski optics on an upright Zeiss microscope. Caged calcium was released from DM-nitrophen and the number of granules present before and 1 min after opening the UV shutter was directly counted.

Major Findings:

Sea urchin egg cortical granule-plasma membrane docking sites can have, on average, nine active fusion complexes. Even though one active fusion complex can initiate fusion, the presence of multiple fusion complexes has profound effects on the behavior of triggered exocytosis. Calcium-secretion coupling represents a manifestation of 'receptor reserve'. Fractional activation of fusion complexes does not correspond directly to the fractional biological response. Receptor reserve describes the situation where activation of a single receptor may invoke a maximal biological response despite the fact that most receptors have not been activated. The behavior is also observed in calcium-triggered exocytosis and may explain the sharp transition observed in calcium activation curves. The pH activation curves observed in pH triggered viral fusion may also be a manifestation of 'receptor reserve'.

Fusion complexes are randomly distributed between granules as a Poisson process. Granules can have, on average, more than one active fusion complex. Time-dependent and steady-state behavior following calcium activation can be fit by our kinetic model based on the hypothesis that activated fusion complexes are randomly distributed as a Poisson process. The maximal rate and extent of fusion reached a plateau at different calcium concentrations in agreement with this kinetic model. The complex behaviors observed in calcium triggered exocytosis can be explained with our kinetic model.

We have developed and applied a new paradigm for considering the consequences of having different distributions of spare fusion-complexes. Under the assumptions that activated fusion complexes are randomly distributed as a Poisson process and an active fusion complex mediates a fusion event between a granule and the plasma membrane in a probabilistic manner, we can explain many of the calcium-dependent exocytotic behaviors observed in this preparation. These include, 1) the sigmoidal dependence of the extent of fusion on pCa, 2) the increase in fusion rate with increasing calcium concentrations, 3) the difference in the calcium concentration necessary to reach the plateau for the extent of fusion and fusion rate, 4) shifting of the calcium activation curve with inhibitors of exocytosis, and 5) aging, where a progressive decrease in calcium sensitivity is observed with time in the planar cortex. This model succeeded in describing calcium-dependent exocytotic behaviors in the sea urchin with the simple hypothesis that an increase in calcium corresponds to an increase in the average number and perhaps efficacy of fusion complexes at docking sites.

The relationship between docking/activation and subsequent calcium triggered fusion was examined by altering the pH of solutions matched for free calcium and developing a quench protocol based on the rapid removal of calcium during the fusion process. No detectable changes in the calcium activation curves were observed following a change in solution pH from 6.7 to 7.4 and matched for free calcium. Reliable kinetic data can be obtained following rapid switching between solutions. The fusion process can be quenched even during the fast decay associated with high calcium concentrations.

Significance to Biomedical Research and the Program of the ICD:

The major significance of this study is that we now have a theoretical model which can explain the complex biological behavior observed during triggered exocytosis. Human maladies of secretion, such as diabetes, infertility, and neuronal and endocrine dysfunction, may now be tested for lesions in the number, distribution, and efficacy of fusion complexes at membrane fusion sites.

Proposed Course:

The dependence of fusion complex efficacy on calcium concentration will be characterized. The effects on the rate and extent of calcium triggered exocytosis due to alterations in physical parameters of the system (lipid composition and temperature for example) will be evaluated using the kinetic model. The relationships between the calcium activation step and the subsequent fusion kinetics will be studied using a rapid quench protocol and evaluated using the kinetic model. The results from local and kinetically specific assays for docking, activation, and fusion including flash photolysis, dynamic light scattering, and polarization will be evaluated using the theoretical model.

Publications:

None

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER
 Z01 HD 01416-03 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Tissue Imaging in Cell Biology

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	L. Margolis	Visiting Scientist	LTPB, NICHD
Others:	S. Glushakova	Visiting Scientist	LTPB, NICHD
	B. Baibakov	Special Volunteer	LTPB, NICHD
	W. Fitzgerald	Special Volunteer	NASA
	N. Amichay	Special Volunteer	NASA

COOPERATING UNITS (if any)

NASA (N. Pellis)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Intercellular Interactions

INSTITUTE AND LOCATION

NIH, NICHD, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

2

PROFESSIONAL:

2

OTHER:

2

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☒ (b) Human tissues ☐ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Continuing to develop imaging and culturing techniques for three-dimensional multicellular systems, we have grown live blocks of human and mouse lymphoid tissue in a rotating wall vessel bioreactor(RWV). Fragments of lymphoid tissue remain viable in the RWV for at least two weeks. Imaging of individual cells with confocal microscopy reveal cell migration from the tissue into the media and back as it occurs in vivo where peripheral blood cells migrate through the lymphoid organs. We have developed the means to control cell migration by embedding the tissue fragments in collagen or agarose gels. Embedding does not impair tissue viability: the gel-embedded blocks of tissue consume oxygen and glucose, produce IgG into the media similar to the controls and retain the basic elements of tissue architecture, including B cell-rich germinal centers surrounded by T cells. Three-dimensional reconstruction of confocal images of embedded blocks of tissue cultured for 2 weeks in RWV reveal a network of follicular-dendritic cells in germinal centers, a key element of antigen-presenting system. Blocks of tissues in RWV are productively infected with HIV-1 providing a unique in vitro model for studying HIV pathogenesis by simulating the basic elements of human lymphoid tissue in vivo. This model allows us to address basic questions of migration, homing and kinetics of normal and HIV-infected cells. In the absence of animal models for HIV infection leading to AIDS, the developed model provides a unique system to study pathogenesis in human lymphoid tissue.

Project Description:Objectives:

One of the major challenges of modern cell biology is to understand the effect of cell microenvironment on cell functioning. Cells interact with each other and with non-cellular substrates. These interactions determine how cells function within tissues. Such kinds of complex interactions are the basis for all aspects of normal and pathological immune responses of lymphoid tissues in vivo. Out of the context of real tissue, even when many types of cells are co-cultured together in cell monolayers or in cell suspensions, cell behavior becomes significantly different.

Important reasons as to why cell behavior continues to be studied in models which poorly represent the situation in vivo are (i) the lack of the methods to support cell viability in three-dimensional biological microenvironment for long enough to study cellular processes and (ii) difficulties of imaging cells in this microenvironment. To bridge the gap between real tissues in vivo and their in vitro models, we are developing new systems of culturing three-dimensional lymphoid tissues. This task becomes even more important when HIV infection is concerned, since HIV pathogenesis involves complex interactions between lymphoid tissue and peripheral blood as there is no adequate animal model for HIV infection. In the present stage of the project we have attempted to culture and to perform imaging analysis on blocks of human tonsils, mouse and human lymph nodes and mouse thymus in the rotating wall vessels.

Methods Employed:

Preparation of tissue blocks. Mouse thymus tissue and mouse lymph nodes are obtained from freshly killed mice. Human lymph nodes are obtained as surgical specimens from deceased persons in sterile saline from NDRI. The human tonsils are obtained as surgical specimens from tonsillectomy patients in sterile saline from Children's Hospital (DC). Cauterized tissue, badly infected tissue, or any other undesirable tissue is dissected away. The remaining tissues are carefully dissected and sliced into pieces of approximately 1x1x1 mm. Collagen embedding is performed using rat tail collagen type 1 isolated from rat tail tendons. The collagen is polymerized by raising the pH and temperature. Agarose embedding employs the use of InCert agarose. The temperature is decreased to polymerize the agarose. The tissue is embedded in a sandwich procedure in 8 well or 16 well chamber slides. When polymerization is complete, the chamber slides are separated to release the gel-covered tissue blocks.

Culturing in RWV. The tissue pieces are placed inside the RWV and the vessel is filled with culture medium. In embedding experiments, RWVs of nonembedded, collagen embedded, and agarose embedded tissue pieces are run side by side with the same number of equivalent sized pieces in each vessel. The RWV is placed on a base and rotated at a speed of 30 - 50 rpm depending on the type and size of tissue blocks. O₂ readings are gathered from the blood gas

readings performed throughout the culture length. Glucose readings are taken before and after each refeeding of the cultures so that glucose consumption can be calculated.

Analysis of the cultured tissue. Medium samples are pulled from cultures and assayed for human IgG using the antibody sandwich ELISA technique. Pieces of tissue are collected and fixed in 4.5% formaldehyde from culture vessels at day 0, day 7, and day 14. For detection of apoptotic cells, paraffin histological sections were stained with the Oncor Apoptag In Situ Apoptosis Detection Kit, which uses direct immunoperoxidase detection of digoxigenin-labeled genomic DNA. Sections were immunostained for T cells, B cell, follicular dendritic cells using monoclonal mouse antibodies (DAKO, Carpinteria, CA) against CD3, CD20, CD21 correspondingly. Either alkaline phosphatase- or peroxidase-labelled or FITC-labelled goat anti-mouse antibodies (Boehringer Mannheim, Indianapolis, IN) were used as secondary antibodies. No immunostaining was observed in these control experiments.

HIV infection and virus production. To infect histocultures we used a laboratory strain of HIV-1, LAV.04/A3.01 obtained through the AIDS Research and Reagent Program, Division of AIDS, NIAID, NIH. To inoculate tonsil histocultures, 1-10 μ l of clarified virus-containing medium was applied onto each tissue block. The multiplicity of infection varied from 400 to 900 TCID₅₀ (50% tissue culture infective dose) per block. 24 hrs post infection, the blocks of tissue were embedded in agarose as described above. The concentration of p24 in medium was measured by HIV-1 p24 antigen ELISA ("Retro-Tek", Cellular Products Inc., Buffalo, NY).

Major Findings:

1. Fragments of human lymphoid tissue cultured in the RWV remain viable for at least two weeks; however, gradual cell depletion is observed. Cells from the tissue migrate into the surrounding medium. Migration of cells out of the tissue is reversible: labelled cells from the peripheral medium re-populate the tissue and their fate can be followed by imaging with confocal microscopy. Thus a two-compartment model of the lymphoid system consisting of blocks of solid lymphoid tissue interacting with single peripheral lymphocytes in fluid medium has been developed.
2. Migration of cells may be partially prevented by embedding the tissue fragments in collagen gels. Cell migration is completely stopped when the tissues are embedded in agarose gels. Thus the exchange rate of cells in the periphery and in the lymphoid tissue in a two-compartment model of lymphoid tissue in vitro can be modulated.
3. Agarose-embedded blocks of tissue cultured in the RWV remain alive for two weeks. They consume oxygen and glucose equivalent to control non-embedded tissues. The consumption of agarose-coated tissue produces IgG in the medium. They retain the basic elements of their architecture including aggregates of T cells surrounding B cell-rich germinal centers. Imaging of the specimens with confocal microscopy together with 3-dimensional reconstruction techniques

demonstrate an elaborate network of follicular-dendritic cells in germinal centers. Agarose-embedded blocks of tissue retain more cells and reveal less apoptotic nuclei than non-embedded controls. Thus a new method of long-term culturing of lymphoid tissue has been developed.

4. Agarose-embedded blocks of tissue are productively infected with HIV-1 (LAV) as measured by increases of p-24 and infectious virus in the medium and HIV RNA-positive cells inside the tissue. Thus HIV-pathogenesis can be studied in a new in vitro model where the structure of lymphoid tissue and interactions with the periphery are preserved.

Significance to Biomedical Research and the Program of the ICD:

The ability to keep lymphoid tissue which preserves its architecture and immunological functions under controlled conditions in vitro in three dimensions is important for answering basic questions about the dynamic structure of the tissue and for studying the mechanisms of its functioning. Culturing of lymphoid tissue and peripheral cells in the Rotating Wall Vessel Bioreactor allows us to simulate the high level of complexity of the immune system and to address the questions of lymphocyte migration and homing. In the absence of animal models for various immune disorders, including HIV infection leading to AIDS, the model developed provides a unique system to study pathogenesis in human lymphoid tissue.

Propose Course:

1. Further characterize the developed model of lymphoid tissue in vitro.
 - a. To study kinetics of cell trafficking through the lymphoid tissue.
 - b. To study the fate and homeostasis of lymphocytes entering the lymphoid tissue using various reporting dyes in combination with analytical confocal microscopy
 - c. To study the ability of the system to respond by production of specific antibodies and cytotoxic lymphocytes to foreign antigens.
2. To test the efficiency, kinetics and pathogenesis of the HIV infection in uni- and two-compartment cultures in the RWV.
 - a. To infect separately peripheral cells and lymphoid tissue and co-culture with the non-infected counterpart.
 - b. To simulate the in vivo kinetics of lymphocyte replenishment and to study the effects of the rate of the replenishment on the HIV infection.
 - c. To analyze cell subsets in the HIV-infected lymphoid system in vitro.

Publications:

Margolis L, Baibakov B, Collin C, Simon SA. Dye-coupling in three-dimensional histoculture of rat lingual frenulum, *In Vitro Cell Dev Biol-Animal* 1995;31:456-461.

Margolis LB, Glushakova SE, Baibakov BA, Collin C, Zimmerberg J. Confocal microscopy of cells implanted into tissue blocks: Cell migration in long-term histocultures, *In Vitro Cell Dev Biol* 1995;221-226.

Margolis LB, Glushakova S, Baibakov B, Zimmerberg J. Syncytium formation in cultured human lymphoid tissue: Fusion of implanted HIV glycoprotein 120/41-expressing cells with native CD4⁺ cells, *AIDS Res Hum Retro* 1995;11:697-704.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01417-02 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Energetics of the Interaction Between Water, Membranes, and Macromolecules

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: A.L. Yervey Head SMAMS, LTPB, NICHD

Others: Y. Xu IRTA Fellow SMAMS, LTPB, NICHD

COOPERATING UNITS (if any)

None

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Metabolic Analysis and Mass Spectrometry

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither

☐ (a1) Minors

☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Preliminary results from the single stage pressure reduction electrospray ionization instrument (API-1) show that very large water clusters, containing more than 25 water molecules and a single proton can be observed. These clusters can be reliably observed, albeit not demononstrably from equilibrium conditions as yet. The reliability can, however, be used to calibrate the instrument's mass scale. Such an approach to calibration could have significant advantages to analytical chemists because it permits mass scale calibration without introduction of contaminating molecules into the instrument. Similar large clusters containing 15-20 water molecules have been obtained with the tetraphenyl arsonium ion, although not yet from equilibrium conditions.

Project DescriptionObjectives:

The purpose of this work is the characterization of the energy required for non-covalent bonds to exist between molecules or ions in solution. Knowledge of the the energetic requirements of bond formation is fundamental to an understanding of molecular interactions. Much of the current activity in theoretical studies of molecular configuration is utterly dependent upon the values assigned to solute/solvent or peptide/metal bonds, yet reliable quantities often do not exist. We are using adaptations of mass spectrometric concepts that have been applied to determinations of energetics of cluster formation in systems of small ions and molecules to determine the energetics of formation in several classes of biologically significant interactions. The bonds of particular interest for this work are those involved in interactions of solutes, including lipids amino acids and small model peptides, with water.

The interactions between water and solute ions will lead to a determination of the enthalpy and free energy of solvation. The initial focus of this portion of the project is the determination of the enthalpy of solvation of $(C_6H_5)_4As^+$ and $(C_6H_5)_4B^-$ independent of any extra-thermodynamic assumptions. Measurements of conductance and dipole potentials for lipid bilayers are dependent upon values of the Gibbs Free Energy of solvation, ΔG_{solv} . At the present, the standard ions $(C_6H_6)_4As^+$ and $(C_6H_6)_4B^-$ (F_4As^+ , F_4B^-) are assumed to have equal values of ΔG_{solv} and ΔH_{solv} , the enthalpy of solvation. This assumption is widely recognized as an unsatisfactory approximation that is used in the absence of the ability to determine energetics of these ions separately. We are in the process of determining the single ion solvation enthalpy of these hydrophobic organic ions using studies of gas phase ion-molecule chemistry under conditions of chemical equilibrium.

Methods Employed:

A series of classic studies done more than 20 years ago by Kebarle, et al, Field et al and recently extended into new systems by Kebarle demonstrate that the energetics of bulk phase solvent-solute interactions can be determined equivalently in the gas phase when the proper conditions are obtained. The early work showed that, under equilibrium conditions, the energetics of sequential additions of solvent molecules to a gas phase ion are

equivalent to the best values of bulk phase energetics when a sufficient number of solvent molecules are present in a cluster. That is, for gas phase cluster ions of solute and solvent, the free energy and enthalpy of solvation are equivalent to the best determination of those values determined in the bulk phase. The significance of this equivalence is, however, that the values determined by the gas phase measurements are independent of any extra-thermodynamic assumptions. In other words, the equilibrium gas phase measurements potentially represent a better approximation to the thermodynamic ideal value than the bulk phase measurements do.

We are using an adaptation of the approach used in Kebarle's determinations of the free energy and enthalpy of solvation of alkali and halide ions. The critical portion of the work is the attainment of equilibrium conditions both in the "high pressure" ion formation portion of the mass spectrometer and in the low pressure portion of the ion optics proximal to the ion formation region. These equilibrium states must be achieved in combination with stable conditions for the formation of the ion clusters of interest. In our case that requires the introduction of the non-volatile organic solutes into the vapor state under well defined conditions of temperature, pressure and solvent concentration. The experimental challenge is in the degree of refinement required for the achievement of conditions necessary for equilibrium to be attained.

The initial experimental approach tried was an adaptation of an existing thermospray ionization source. While initial results from this approach were promising, we have concluded that conditions consistent with equilibrium ion distributions lead to unsatisfactorily low cluster intensities. The second approach that was investigated was to adapt a capillary interface electrospray mass spectrometer to yield cluster ions with an intensity distribution that is unperturbed from the equilibrium state in which they are formed. This too has been abandoned due to the consequences of the perturbations in ion intensity distribution during the free jet expansion at the capillary's termination. The present approach being developed is an adaptation of a reduced pressure electrospray ion source in conjunction with a single stage pressure reduction to mass analysis using a cryopumped mass analyzer. This approach has recently been shown by Kebarle to yield equilibrium ion cluster distributions and to reproduce alkali ion ΔH_{Solv} values obtained under other conditions. This combination of electrospray ionization of non-volatile organic molecules coupled with careful control of water vapor content in the cluster formation volume is virtually certain to yield the desired equilibrium ion cluster distributions.

Major Findings:

1. Proton Enthalpy of Solvation

Measurement of the enthalpy change for the successive additions of water molecules to a hydrated proton cluster ion are the archetypical reaction for studies of hydration. This reaction has been studied extensively in the past, and is being used as the model and proof of method in the present work. Since the hydrophobic quaternary ions initially of interest can be introduced to the ion source from a flowing water stream, use of the water cluster system as the starting point for measurements is a reasonable approach.

Preliminary results from the single stage pressure reduction electrospray ionization instrument (API-1) show that very large water clusters, containing more than 25 water molecules and a single proton can be observed. These clusters can be reliably observed, albeit not demononstrably from equilibrium conditions as yet. The reliability can, however, be used to calibrate the instrument's mass scale. Such an approach to calibration could have significant advantages to analytical chemists because it permits mass scale calibration without introduction of contaminating molecules into the instruemnt. Similar large clusters containing 15-20 water molecules have been obtained with the tetraphenyl arsonium ion, although not from equilibrium conditions as yet.

2. Electron Multiplier Response

A fundamental aspect of using mass spectrometric intensities for quantitative analysis is a knowledge of the number of electrons produced by the impact of an ion on the first stage of an electron multiplier. Heavier ions are known to produce fewer ions than lighter ones, and multiply charged ions of a given mass produce a higher electron yield than singly charged ions. Observed ion intensities can be corrected to true values if the electron yield function is known. Friedman and Xu have previously quantified this phenomenon for high energy ($>1\text{MeV}$) ions, but knowledge of the detailed behavior of electron yields at more conventional multiplier voltages, ca 3kV, have been lacking. We have recently completed a study that provides the functional dependence of electron yield as a function of mass and charge state for multipliers operating in the 3kV range. Use of this function for correcting cluster ion intensity distributions as well as the more generally applicable charge state intensity distribution for multiply charged electrospray ions will play an important role in our future quantitative measurements.

Significance to Biomedical Research and the Programs of the ICD:

The work undertaken in this project represents studies that are fundamental to issues of physical aspects in structural biology. The initial studies will lead to the resolution of a significant controversy in the determination of membrane dipole potentials. Having developed the methodology for this particular study, these techniques will be used for studies of the energetics of lipid, amino acid and peptide hydration. The larger significance of this work is associated with these latter measurements. Knowledge of the hydration energetics will be of fundamental importance in estimations of the energetics of membrane self assembly and in determinations of energetics of protein folding by computational biology.

Proposed Course:

1) Complete ion source modifications to API-1 single stage pressure reduction instrument. This will permit obtaining stable ion cluster ratios as a function of water partial pressure. Develop systematic investigation of tetraphenyl quaternary ions: F_4N^+ , F_4P^+ , F_4As^+ as well as related anions.

Publications:

Xu Y, Zhang X, Yergey AL. Studies of transition metal ion complexes with amino acids and peptides by electrospray ionization mass spectrometry. J Am Soc Mass Spectrom, 1995, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01418-02 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Regulation of Mitosis at the Cellular Level

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: F. Suprynowicz Senior Staff Fellow LTPB, NICHD

Others: T. Whalley Visiting Fellow LTPB, NICHD

A. Sokoloff Visiting Scientist LTPB, NICHD

COOPERATING UNITS (if any)

Laboratory of Biochemistry, NHLBI (S.-K. Kong, Ph.D.); Dept. of Physiological Sciences, Univ. of Newcastle, U.K. (L. Groigno, Ph.D.)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Membrane and Cellular Biophysics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither

☐ (a1) Minors

☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The onset and completion of mitosis are universally controlled by activation and inactivation of the p34^{cdc2} protein kinase. Molecular studies have shown that p34^{cdc2} kinase activity is regulated by phosphorylation at multiple sites, and by means of a physical association with B-type cyclins. Physiological studies have demonstrated that cytoplasmic Ca²⁺ levels increase transiently at specific stages of the cell cycle, and that manipulations of cytoplasmic Ca²⁺ can accelerate or delay entry and exit from mitosis. To develop a more complete understanding of mitotic regulation at the cellular level, we have focused on defining the mechanism by which cytoplasmic Ca²⁺ transients may control p34^{cdc2} activation and inactivation. We previously established a link between Ca²⁺ and p34^{cdc2} by directly demonstrating that physiological levels of Ca²⁺ induce premature inactivation of the p34^{cdc2} kinase in permeabilized sea urchin embryos. We now show that activators of protein kinase C block p34^{cdc2} activation and inactivation in vivo by suppressing cytoplasmic Ca²⁺ transients which are required for the dephosphorylation of p34^{cdc2} at key regulatory sites. Our findings suggest a model for mitotic regulation in the sea urchin embryo whereby the association of p34^{cdc2} with cyclin B induces phosphorylation of p34^{cdc2} at activating and inhibitory sites, while cytoplasmic Ca²⁺ transients trigger the dephosphorylation of these sites that directly controls p34^{cdc2} protein kinase activity.

Project Description:Objectives:

The onset and completion of mitosis are universally controlled by activation and inactivation of the p34^{cdc2} protein kinase. Molecular studies have shown that p34^{cdc2} kinase activity is regulated by phosphorylation at multiple sites, and by means of a physical association with B-type cyclins, proteins which oscillate in abundance over the course of the cell cycle. Physiological studies have demonstrated that cytoplasmic Ca²⁺ levels increase transiently at specific stages of the cell cycle, and that manipulations of cytoplasmic Ca²⁺ can accelerate or delay entry and exit from mitosis. To develop a more complete understanding of mitotic regulation at the cellular level, we have adopted an approach that integrates techniques of biochemistry and cell physiology using permeabilized cells and cell-free assays. In particular, we have focused on defining the mechanism by which cytoplasmic Ca²⁺ transients may control p34^{cdc2} activation and inactivation.

Methods Employed:

Early embryos of the sea urchin *Lytechinus pictus* are ideal for these studies, as they exhibit relatively rapid cell cycles (1-2h) with a high degree of mitotic synchrony. Quantities of cells suitable for biochemical analysis are easily obtained, and the large size of the embryos (110 μ m diameter) greatly facilitates electroporation and microinjection.

Intracellular Ca²⁺ measurements - The concentration of free cytoplasmic Ca²⁺ was monitored over the course of the cell cycle in embryos injected with the Ca²⁺ indicator, Fura-2-dextran, using confocal microscopy and fluorescence ratio imaging.

Kinase assays - Embryos were lysed using ultrasonic disruption in a medium that reproduces characteristics of the intracellular environment. Protein kinase C (PKC) and p34^{cdc2} protein kinase activities were determined using a phosphocellulose filter binding assay. PKC activity was measured by radioactive phosphorylation of the selective peptide substrate [ser²⁵]PKC(19-31) in the presence of Ca²⁺, phorbol 12-myristate 13-acetate (PMA) and phosphatidyl serine. p34^{cdc2} kinase activity was assayed in the absence of Ca²⁺, using histone H1 as the substrate.

Cyclin B quantitation - Cyclin B levels were determined using SDS polyacrylamide gel electrophoresis (SDSPAGE) and phosphor-imager analysis of embryos labeled with [³⁵S] methionine. The association of cyclin B with p34^{cdc2} was determined by co-adsorption on immobilized p13^{suc1}.

Phosphotyrosine immunoblots - In many systems, the protein kinase activity of p34^{cdc2} is negatively regulated by phosphorylation at Tyr15. We studied the effect of PKC agonists on the

phosphorylation state of this site using anti-phosphotyrosine immunoblots. p34^{cdc2} was isolated by adsorption to immobilized p13^{suc1} and fractionation on SDSPAGE. Gels were electrophoretically transferred to nitrocellulose and incubated with an anti-phosphotyrosine monoclonal antibody, followed by [¹²⁵I]-labeled protein A.

Phosphoamino acid analysis - The protein kinase activity of p34^{cdc2} is known to be positively regulated by phosphorylation at Thr161. To study the effect of PKC agonists on the phosphothreonine content of p34^{cdc2}, proteins were labeled to steady state levels in vivo with [³²P]orthophosphate. Labeled p34^{cdc2} was adsorbed to immobilized p13^{suc1}, analyzed by SDSPAGE, transferred to Immobilon PVDF membranes and visualized by autoradiography. Using the autoradiograph as a template, the p34^{cdc2} band was excised and incubated in 6N HCl to hydrolyze peptide bonds. Phosphoserine, phosphothreonine and phosphotyrosine were separated by thin-layer chromatography on cellulose plates and visualized by autoradiography.

Phosphopeptide mapping - To determine whether or not cell cycle-related oscillations in the phosphothreonine content of p34^{cdc2} were due to phosphorylation/dephosphorylation at Thr161, Immobilon transfers of ³²P-labeled p34^{cdc2} (see above) were digested with trypsin and subjected to two-dimensional analysis on thin-layer cellulose plates involving electrophoresis at pH 1.9 and ascending chromatography in the second dimension.

Major Findings:

We previously established a link between intracellular Ca²⁺ transients and p34^{cdc2} by directly demonstrating that physiological levels of Ca²⁺ induce premature inactivation of the p34^{cdc2} kinase in permeabilized sea urchin embryos. Since no effect of Ca²⁺ on p34^{cdc2} kinase activity was observed in more dilute homogenates, it is clear that p34^{cdc2} is not intrinsically sensitive to Ca²⁺, but that one or more co-factors are involved in the inactivation pathway. We have investigated whether or not a Ca²⁺-dependent form of PKC may be such a co-factor, as numerous reports have documented effects of PKC activators on cell cycle progression.

PKC activators block activation and inactivation of p34^{cdc2} -

1) We have shown that activators of PKC, such as PMA and 1,2-dioctanoylglycerol, arrest the sea urchin cell cycle in a stage-dependent manner. Administration of these agonists prior to the beginning of S phase prevents entry to mitosis (arrest at the G2/M boundary), while administration during prophase blocks exit from mitosis (arrest at metaphase). In contrast, 4a-PMA, a related compound that does not activate PKC, has no effect on cell cycle progression. The action of PMA is not to generally suppress cell metabolism, as rates of protein and DNA synthesis are not affected. PMA-induced cell cycle arrest at G2/M has been reported previously, however this is the first demonstration that PMA blocks exit from mitosis.

2) We have established that PKC activation prevents activation and inactivation of the p34^{cdc2} protein kinase. p34^{cdc2} is not activated in PMA-treated embryos which are arrested at G2/M, whereas p34^{cdc2} is not inactivated in embryos which are arrested at metaphase. Elevated protein kinase activity is responsible for the metaphase arrest, as it is reversed by addition of the protein kinase inhibitor, staurosporine. These results indicate that inappropriate activation of PKC arrests the embryonic cell cycle by blocking activation and inactivation of the key regulatory protein, p34^{cdc2}. PKC itself is not essential for cell cycle progression, since embryos injected with the highly selective PKC inhibitor peptide, PKC(19-31), undergo multiple cell divisions.

PKC activation uncouples cycles of cyclin abundance and p34^{cdc2} phosphorylation - The association of cyclin B with p34^{cdc2} during interphase promotes the phosphorylation of p34^{cdc2} at both inhibitory (Tyr15) and activating (Thr161) sites. The dephosphorylation of Tyr15 during late G2 phase (by the cdc25 phosphatase) turns on p34^{cdc2} protein kinase activity, while the dephosphorylation of Thr161 during anaphase is required for kinase inactivation.

1) We have shown that the activation of PKC prior to S phase (resulting in G2/M arrest) does not inhibit the association of p34^{cdc2} with cyclin B nor the phosphorylation of p34^{cdc2} on tyrosine residue(s). However, dephosphorylation of this site(s) does not occur, which explains why p34^{cdc2} is not activated in these embryos. The results suggest that PKC activation in some way inhibits the cdc25 phosphatase pathway.

2) We have shown that the activation of PKC during prophase (resulting in metaphase arrest) does not prevent the destruction of cyclin B, which occurs at the same time as in control embryos. However, the dephosphorylation of p34^{cdc2} at a site identified as Thr161 is inhibited, which explains why p34^{cdc2} is not inactivated in these embryos. It is widely thought that p34^{cdc2} kinase inactivation during mitosis is the direct result of cyclin B proteolysis. Our finding that PMA blocks p34^{cdc2} inactivation without preventing the destruction of cyclin B demonstrates that these two processes can be uncoupled, and stresses the importance of studying the pathway controlling Thr161 dephosphorylation.

PKC blocks p34^{cdc2} activation/inactivation by suppressing cytoplasmic Ca²⁺ transients - In a number of cells, the activation of PKC suppresses transient increases in the concentration of free cytoplasmic Ca²⁺. Since activators of PKC arrest the sea urchin cell cycle at stages identical to those observed following microinjection of Ca²⁺ chelators, we have tested the hypothesis that PKC activation suppresses cell cycle-related cytoplasmic Ca²⁺ transients that have been described in this system.

(1) Using fluorescence ratio imaging of embryos injected with a Ca²⁺ indicator, we have directly demonstrated that PMA treatment eliminates cytoplasmic Ca²⁺ transients. Typical cell cycle-related Ca²⁺ fluctuations are detected in control embryos not treated with PMA.

2) We have obtained evidence that the suppression of cytoplasmic Ca^{2+} transients by PMA inhibits $\text{p34}^{\text{cdc}2}$ kinase activation and inactivation. In embryos arrested at G2/M with a low level of $\text{p34}^{\text{cdc}2}$ kinase activity, experimental induction of a cytoplasmic Ca^{2+} transient (treatment with ionomycin or Ca^{2+} ionophore A23187) triggers $\text{p34}^{\text{cdc}2}$ activation. In contrast, embryos arrested in mitosis with a high level of $\text{p34}^{\text{cdc}2}$ kinase activity undergo $\text{p34}^{\text{cdc}2}$ inactivation following induction of a similar Ca^{2+} transient.

3) Our previous finding that elevated cytoplasmic Ca^{2+} inactivates $\text{p34}^{\text{cdc}2}$ in permeabilized sea urchin embryos in the absence of cyclin B proteolysis suggested the hypothesis that a Ca^{2+} -dependent mechanism initiates $\text{p34}^{\text{cdc}2}$ inactivation, and that the destruction of cyclin B serves to ensure that another round of kinase activation does not take place before another round of DNA replication. This hypothesis is supported by our present finding that activators of PKC block $\text{p34}^{\text{cdc}2}$ inactivation in vivo by preventing the increase in cytoplasmic Ca^{2+} that normally takes place at anaphase onset, rather than by interfering with the destruction of cyclin B. Together, these results suggest a model for mitotic regulation in the sea urchin embryo whereby the association of $\text{p34}^{\text{cdc}2}$ with cyclin B induces phosphorylation of $\text{p34}^{\text{cdc}2}$ at activating and inhibitory sites, while cytoplasmic Ca^{2+} transients trigger the dephosphorylation of these sites that directly controls $\text{p34}^{\text{cdc}2}$ protein kinase activity.

Significance to Biomedical Research and the Program of the ICD:

Cell division is a fundamental process underlying all aspects of growth and differentiation. An understanding of the control of cell division is essential if we are to understand the basis of early embryonic development and cell differentiation. In particular, an appreciation of the mechanisms controlling the driving force for mitosis and meiosis, that is the $\text{p34}^{\text{cdc}2}$ protein kinase, should lead us to new approaches for therapeutic intervention in diseases characterized by dysfunctional cell division, such as infertility and cancer.

Proposed Course:

1) A variety of evidence supports the idea that protein phosphatase-1 (PP-1) is responsible for the dephosphorylation of $\text{p34}^{\text{cdc}2}$ at Thr161 which abolishes $\text{p34}^{\text{cdc}2}$ protein kinase activity at anaphase onset. The manner in which PP-1 is regulated is not fully understood, however its activity is known to be modulated by reversible association with the heat-stable protein, phosphatase inhibitor-1 (I-1). The association of I-1 with PP-1 in turn is controlled by the phosphorylation state of I-1. Our evidence that elevated cytoplasmic Ca^{2+} triggers inactivation of $\text{p34}^{\text{cdc}2}$ at anaphase onset is consistent with these observations, since the dephosphorylation (and inactivation) of I-1 is known to be mediated by the Ca^{2+} -dependent phosphatase, calcineurin. We will study the involvement of calcineurin in $\text{p34}^{\text{cdc}2}$ inactivation by microinjecting a selective peptide inhibitor of calcineurin into mitotic sea urchin embryos, and by directly investigating the effect of this peptide on the activities

of PP-1 and p34^{cdc2} in cell-free preparations.

2) During mitotic prophase, a nuclear Ca^{2+} transient immediately precedes nuclear envelope breakdown (NEB) and chromosome condensation. Several studies have demonstrated that this Ca^{2+} transient is both necessary and sufficient to trigger these mitotic events in sea urchin embryos, and that the multifunctional Ca^{2+} and calmodulin-dependent protein kinase (CaM kinase) is required. Since our results indicate that p34^{cdc2} already has been activated at the time of this Ca^{2+} transient, and the results of other laboratories indicate that the p34^{cdc2}-cyclin B complex translocates from the cytoplasm to the nucleus just prior to NEB, we will investigate the possibility that the nuclear Ca^{2+} transient and CaM kinase trigger nuclear translocation of the active p34^{cdc2} complex. Recombinant sea urchin cyclin B (from Dr. Tim Hunt) will be fluorescently labeled and injected into sea urchin embryos during interphase. We may then use fluorescence microscopy to follow the nuclear translocation of this protein and to study the effect of Ca^{2+} chelators and selective CaM kinase peptide inhibitors on the translocation.

Publications:

None

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01419-02 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Localization of Membrane Fusion Proteins

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	J. Zimmerberg	Chief	LTPB, NICHD
Others:	A. Sokoloff	Visiting Associate	LTPB, NICHD
	T. Whalley	Visiting Fellow	LTPB, NICHD

COOPERATING UNITS (if any)

None

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Membrane and Cellular Biophysics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Intracellular membrane fusion is crucial for the secretory and endocytotic pathways. Recently, great advances have been made in identifying proteins which have been implicated in playing crucial roles in these processes. Among these proteins is *N*-ethylmaleimide (NEM)-sensitive fusion protein (NSF) and a whole family of proteins with which it interacts. We have used 2 different preparations in order to investigate the involvement of NSF in intracellular fusion events. These are the Ca^{2+} -stimulated exocytosis of sea urchin egg cortical granules and the GTP-triggered fusion of rat liver endoplasmic reticulum microsomes. Cortical granule exocytosis is an ATP- and cytosol-independent Ca^{2+} -stimulated membrane fusion event which is inhibited by NEM. We have established that cytosol is unable to restore fusion to the isolated exocytotic machinery inactivated by NEM. We have found using immunoblot analysis that sea urchin eggs contain NSF in their cytosol but that this protein is absent from an in vitro preparation which undergoes fusion upon the provision of Ca^{2+} . However, a stable and specific interaction between NSF and the isolated exocytotic machinery can occur under the appropriate conditions. These results suggest that the role of NSF might be in the establishment of a fusion-competent arrangement rather than in membrane fusion per se. Similarly, we found that the sensitivity of the homotypic fusion of endoplasmic reticulum membranes to NEM appeared to have characteristics distinct from those observed in reconstituted membrane transport and associated with NSF. Now, we have shown that additional purification and extraction of endoplasmic reticulum membranes with high salt that increases the GTP-dependent fusion activity of the membranes results in complete removal of the small amounts of NSF detectable in the "crude" membrane preparation. In addition, we found that anti-NSF monoclonal antibodies which are known to functionally inhibit NSF-dependent fusion had no effect on the fusion of endoplasmic reticulum membranes. These data corroborate our previous results and support the conclusion that the NEM-sensitive component required for GTP-dependent fusion of endoplasmic reticulum membranes is distinct from NSF. This suggests a novel, NSF-independent mechanism of fusion.

Project Description

Objectives:

Our primary goal has been to learn more about the proteins which are required for intracellular membrane transport and fusion both during exocytosis and the fusion of endoplasmic reticulum microsomes. In both Ca^{2+} -stimulated sea urchin egg cortical granule exocytosis and the GTP-triggered fusion of rat liver endoplasmic reticulum microsomes the docking step has been accomplished before fusion is triggered experimentally and thus any results pertain to the very last steps - fusion itself. It is to be hoped that by identifying the proteins required for various processes during vesicle transport and membrane fusion we will be able to develop strategies to intervene therapeutically.

Methods Employed:

Preparation of Cell Surface Complex (CSCs)

Sea urchins were purchased from Marinus (Long Beach, CA) and were maintained in aquaria in artificial sea water (435 mM NaCl, 40 mM MgCl_2 , 15 mM MgSO_4 , 11 mM CaCl_2 , 10 mM KCl, 10 mM Hepes, 1 mM EDTA, pH 8.0) at a temperature of 14°C. Eggs were obtained by injecting 0.5 M KCl into the intracoelomic cavity and the jelly coat was removed by passing through 90 μm nylon mesh. The isolated exocytotic machinery (CSCs) was isolated in the following way. Eggs were washed 3 times and resuspended in IM (IM: 220 mM Kglutamate, 500 mM glycine, 10 mM NaCl, 5 mM MgCl_2 , 5 mM EGTA, mM benzamidine, pH 6.7) at 4°C and were homogenized gently until a microscopic examination showed that no intact eggs remained. After homogenization, CSCs were pelleted at 700 x g for 1 minute and resuspended in fresh IM.

Preparation of Egg S100 Cytosol

Eggs were washed into ice cold IM containing 2.5 mM MgATP, 5 mM DTT, 10 $\mu\text{g}/\text{ml}$ aprotinin, pepstatin and leupeptin (cytosol buffer). Eggs were resuspended in an equal volume of cytosol buffer and were gently homogenized by 3 strokes of a loose-fitting teflon pestle in a Potter homogenizer. The homogenate was centrifuged for 1 hour at 4°C and 100,000 x g. The supernatant (S100) was retained, passed through a G25 Sephadex column in IM and frozen at -70°C until use.

Measuring Exocytosis

100 μl aliquots of CSCs were dispensed into flat-bottomed microtiter plates. Exocytosis was triggered by the addition of 100 μl of Ca^{2+} /IM to give the desired Ca^{2+} concentrations. Ca^{2+} concentrations were determined using a calcium-selective electrode and exocytosis was measured in a microtiter plate reader as a decrease in absorbance at a wavelength of 405 nm.

NSF-binding Assay

The binding of NSF to egg cortical fragments was performed in the following way. Planar cortices were isolated by attachment to polylysine-coated plastic petri dishes. They were incubated with cytosol under a variety of conditions. After an appropriate incubation time, the cortices were exhaustively washed and protein detached from the solid support by incubation with 0.1 M Tris, pH 11. Protein was precipitated using trichloroacetic acid and dissolved in sample buffer.

Isolation of endoplasmic reticulum membranes

"Crude" endoplasmic reticulum membranes were isolated by differential centrifugation. Highly purified, salt extraction membranes were prepared by isolating rough endoplasmic reticulum fraction by isopycnic centrifugation followed by extracting membranes with 1M NaCl. Extracted membranes were isolated by isopycnic centrifugation and used in experiments.

Fusion assay for endoplasmic reticulum membranes

For fusion to occur, non-labeled membranes and membranes containing R18 (octadecyl rhodamine B) incorporated into the membrane at a fluorescence self-quenching concentration were brought into contact with polyethyleneglycol (PEG). Fusion was triggered by the addition of GTP and detected by an increase in R18 fluorescence.

Electrophoresis and Western Blotting

Samples were subjected to electrophoresis in 12 % polyacrylamide gels and transferred to PVDF membranes. NSF was detected using the anti-NSF monoclonal antibody 6E6 followed by a rabbit anti-mouse secondary antibody and ¹²⁵I-labelled Protein A.

Major Findings:

1) One of the paradigms for NSF-dependent membrane fusion is that cytosol is able to restore activity to NEM-inactivated fusion systems. We have shown that cytosol fails to restore fusion to the NEM-inactivated isolated exocytotic machinery, even when inhibition is sub-maximal. There is no concentration of NEM whose inhibitory effect is overcome by the addition of cytosol. It has now been firmly established that cytosol has no effect upon the Ca²⁺-sensitivity of cortical granule exocytosis in vitro. 2) We have found that sea urchin eggs contain NSF. This protein is located in the cytosol but not in the isolated exocytotic machinery. 3) Sea urchin egg NSF has the same membrane binding properties that have been found for NSF from other sources. In the presence of non-hydrolyzable analogues of ATP, NSF irreversibly binds to the isolated exocytotic machinery. This binding is prevented by treatment with NEM and recombinant NSF competes for the same binding sites. This binding has no effect on calcium-stimulated exocytosis. These data all suggest that membrane fusion during cortical granule exocytosis is downstream of the actions of NSF. 4)

"Crude" endoplasmic reticulum membranes contain small quantities of NSF as detected by immunoblot analysis. However, the additional purification and extraction of endoplasmic reticulum membranes with high salt which increases the GTP-dependent fusion activity of the membranes results in the complete removal of the small amounts of NSF detectable in the "crude" membrane preparation. We also found that functionally inhibitory anti-NSF monoclonal antibodies had no effect on the fusion of endoplasmic reticulum membranes. These data corroborate the previous results and support the conclusion that the NEM-sensitive component required for GTP-dependent fusion of endoplasmic reticulum membranes is distinct from NSF. This suggests a novel, NSF-independent mechanism of fusion. 5) We synthesized thiol-specific bulky reagents to detect and affinity isolate microsome surface proteins with exposed, functionally important thiol groups. We have raised monoclonal antibodies against these proteins with a view to identifying which play important roles during fusion.

Significance to Biomedical Research and the Program of the ICD:

Intracellular membrane fusion is crucial to the functioning of all cells. It is required for exocytosis, endocytosis, secretory protein trafficking and neurotransmitter release to name but a few. Understanding the mechanisms behind membrane fusion should allow us to develop techniques to intervene in pathological conditions characterized by dysfunctions in these processes.

Proposed Course:

- 1) We shall determine the proteins in the sea urchin exocytotic machinery which interact with NSF using an affinity isolation procedure. This will allow us to develop probes to determine the roles of these proteins in vesicular transport and fusion. This approach can also be used to determine the NSF-binding sites in endoplasmic reticulum membranes.
- 2) We shall use the monoclonal antibodies raised against thiol-containing proteins in order to identify the fusion-related proteins whose thiols are required for fusion.
- 3) We aim to identify the GTP-binding proteins which are required for endoplasmic reticulum membrane fusion.

Publications:

Sokoloff AV, Whalley T, Zimmerberg J. The characterization of N-ethylmaleimide-sensitive thiol groups required for the GTP-dependent fusion of endoplasmic reticulum membranes. *Biochem J*, 1995, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01420-01 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Calcium Triggered Membrane Trafficking in Sea Urchin Eggs

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	S.S. Vogel	Senior Staff Fellow	LTPB, NICHD
Others:	M.S. Cho	Biologist	LTPB, NICHD
	M. Terasaki	Senior Staff Fellow	LN, NINDS
	T. Whalley	Visiting Fellow	LTPB, NICHD
	S. Beushausen	Visiting Scientist	LN, NINDS
	D. Lester	Staff Scientist	FDA

COOPERATING UNITS (if any)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Section on Membrane and Cellular Biophysics

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

1.5

PROFESSIONAL:

1.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Following the fertilization of sea urchin eggs, calcium is released from intracellular stores triggering the fusion of thousands of exocytotic cortical granules with the egg plasma membrane. We have used this preparation to study two aspects of calcium triggered membrane trafficking. 1) We have used light microscopy, confocal microscopy, electron microscopy, and biochemical assays to determine how eggs retrieve excess membrane following cortical granule exocytosis. We find that eggs retrieve excess membrane from the plasma membrane by directly internalizing large 1.5 μ m diameter vesicles. 2) We have developed an assay using an *in vitro* preparation of sea urchin egg plasma membrane and cortical granules to screen drugs for their ability to lyse membranes and/or interfere with calcium triggered exocytosis. We have applied this assay to screen antineoplastic drugs to investigate the possibility that cancer might be treated by interfering with intracellular trafficking (which should be required to maintain neoplastic growth). Our assay detected two drugs which might work by this mechanism. We found that tamoxifen is membrane active and can lyse intracellular vesicles. We found that taxol acts as a reversible inhibitor of cellular secretion, and acts at a step prior to calcium activation of a fusion complex.

Project Description:Objectives:

Our goal is to expand our knowledge of the mechanism of exocytotic membrane fusion and endocytotic membrane fission, and to apply this knowledge to various aspects of pharmacology including drug screening, drug development, and drug delivery. We hope to identify and develop drugs that can modulate these reactions involved in releasing important cellular products into the extracellular milieu and retrieving membrane and extracellular fluids back into the cell. We also believe that if we can understand the cellular mechanisms of transporting macromolecules across biological membranes we may be able to harness that cellular machinery for delivering membrane impermeant therapeutics, such as nucleic acids, to presently inaccessible target sites in cells.

Methods Employed:

To assay endocytic activity, eggs are incubated at room temperature in 100 μ m rhodamine dextran (3,000 MW) and activated with sperm or calcium ionophore. Time points are collected by stopping the reaction on ice and subsequent centrifugation to remove fluorescent dextran from the bulk-solution. Eggs are next lysed and egg-incorporated fluorescence quantitated in a fluorimeter. Endocytosis was also observed in a conventional fluorescence microscope and confocal microscope as fluorescent bodies (1.5 μ m diameter) located inside the egg which do not wash away when dye (rhodamine dextran, RH-414, or FM-143) is removed from the bulk solution.

To screen drugs we prepared a cell surface complex (CSC) from sea urchin eggs. Sea urchins were purchased from Marinus Inc. (Long Beach, CA) Eggs are passed through Nitex mesh and washed three times to remove egg jelly coat. Cortical granules are isolated by first preparing a cell surface complex by homogenizing eggs in PKMEB buffer (50 mM pipes, pH 6.7, 425 mM KCl, 10 mM MgCl₂, 5 mM EGTA, 0.5 mM BAPTA, 5 mM benzamidine) and recovering the low speed pellet composed of CSC's. The calcium dependence of cortical granule-plasma membrane fusion is measured using a microtiter dish adaptation of a turbidity assay described by Sasaki and Epel (1983, Dev. Biol. 98, 327-337) in HENPK buffer (50 mM pipes, pH 6.7, 425 mM KCl, 3.3 mM HEDTA, 3.3 mM EGTA, 3.4 mM NTA, 5 mM benzamidine).

CSC's in HENPK are incubated for 15 minutes at RT with different drugs prior to being aliquoted (100 μ l/point) in 96-well dishes. To trigger fusion, 100 μ l of calcium fortified HENPK is added per point.

Major Finding:

Over the past year we have had 2 major developments which have advanced our knowledge of the mechanism of calcium triggered exocytosis and endocytosis: 1) we have found that cells compensate for the added membrane and lost volume following exocytosis by directly recovering membrane into large endocytic structures, and 2) we have developed an *in vitro* assay for screening drugs and have found that two anti-cancer drugs have profound effects on calcium triggered exocytosis in sea urchin eggs.

Exocytosis in many cell types is often followed by endocytosis. Using various forms of microscopy we have directly shown that large internal structures are rapidly formed *de novo* following exocytosis. This is the first time these large structures have been seen to form subsequent to exocytosis in an intact cell, and these structures are not produced by a treatment which stimulates clathrin-dependent endocytosis. This is a novel form of endocytosis which may represent a new pathway in regulated membrane traffic. Whether triggered by a transient elevation of intracellular calcium as at fertilization or a sustained elevation (using the ionophore, A23187), the kinetics and extent of retrieval were the same, suggesting that the role of intracellular calcium in triggering endocytosis is only to trigger exocytosis. Our experiments have not proven that cortical granule exocytosis is a prerequisite for the rapid retrieval of membrane. Still, the fact that we have never observed retrieval in the absence of cortical granule exocytosis strongly suggests that this is the case.

We have developed a simple *in vitro* assay to screen drugs for their ability to perturb calcium triggered exocytosis. This assay uses membrane components isolated from sea urchin eggs, and has the potential of monitoring membrane effects at the level of the plasma membrane and the secretory granule membrane. Because eggs can be collected from animals without sacrifice, an advantage of this assay is that it does not require the loss of animal life. We screened 18 different drugs with our assay and found that tamoxifen decreased the turbidity of cell surface complexes in calcium free buffers. Tamoxifen is an amphipathic compound which is known to partition into membranes, subsequently altering biophysical properties of the bilayer. This suggests that tamoxifen can

interact with membranes of intracellular organelles, a proposed intracellular receptor site. Interestingly other hydrophobic drugs tested (e.g. daunorubicin) did not lyse granules, suggesting that these compounds may not perturb intracellular membranes. This observation suggests that some of the adverse effects of tamoxifen chemotherapy might be related to this membrane lytic activity. A second potential application of this assay, the identification of drugs which perturb the secretory process, was demonstrated upon treatment of the cell surface complex preparation with taxol. Taxol inhibited calcium-dependent exocytosis, resulting in a requirement for higher calcium. Taxol is known to stabilize microtubules in sea urchin eggs and in mammalian cells, so our finding is consistent with its site of action being the cytoskeleton. Our finding also supports other studies that show that taxol can inhibit secretion from mammalian systems. It is interesting that vinblastine, a drug which destabilizes microtubules, did not perturb calcium triggered exocytosis. Consistent with this, other cytoskeletal modifying agents fail to modulate cortical granule exocytosis. This might reflect the important feature that in unfertilized sea urchin eggs cortical granules are already docked to the plasma membrane and fusion depends on the microtubule network being in a disassembled state. Alternatively, taxol inhibition of calcium triggered exocytosis might be working through a microtubule-independent mechanism. Our finding, in conjunction with reports of taxol inhibition of secretion in many different cell types raises the intriguing possibility that the antineoplastic mechanism of taxol may involve the inhibition of membrane trafficking required for rapid cell growth.

Significance to Biomedical Research and the Program of the ICD:

Exocytosis and the subsequent endocytosis are steps involved in fertilization, synaptic transmission, and hormone release. Understanding the mechanism of exocytosis and endocytosis is important for our understanding of such diverse human maladies as infertility, neuronal and endocrine dysfunction, and diabetes. Our drug screening assay has been developed to the point where it can now be used to identify compounds which modulate intracellular trafficking, and thus may be useful for treating cancer. Our assay can also be used to screen compounds for their potential to treat other diseases which also might involve membrane trafficking.

Proposed Course:

1) Drs. S. Vogel and S. Beushausen will characterize dynmamin, a protein thought to be essential for endocytosis, in sea urchin eggs and study dynamins role in the endocytotic retrieval of plasma membrane following cortical granule exocytosis.

2) Drs. S. Vogel and D. Lester of the FDA will: study how tamoxifen lyses membranes using fast-Fourier infa-red spectroscopy, and b) determine if taxol is affecting granule docking.

Publications:

Vogel SS, Beushausen S, Lester DS. Application of a membrane fusion assay for rapid drug screening, Pharmaceut Res, 1995, in press.

Whalley T, Terasaki M, Cho M-S, Vogel SS. Direct membrane retrieval into large vesicles following exocytosis in sea urchin eggs, J Cell Biol, 1995, in press.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01501-04 LTPB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Role of Lipids in Membrane Rearrangements

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI:	L. Chernomordik	Visiting Scientist	LTPB, NICHD
Others:	A. Chanturiya	Visiting Associate	LTPB, NICHD
	E. Leikina	Visiting Associate	LTPB, NICHD
	M. Cho	Biologist	LTPB, NICHD
	T. Rarick	Medical Staff Fellow	LTPB, NICHD
	J. Zimmerberg	Chief	LTPB, NICHD

COOPERATING UNITS (if any)

LAB/BRANCH

Laboratory of Theoretical and Physical Biology

SECTION

Lipids in Membrane Fusion

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

4

PROFESSIONAL:

3.5

OTHER:

0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☒ (b) Human tissues ☐ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

To elucidate the molecular mechanisms of biological fusion we have studied the lipid-involving fusion intermediates. Membrane fusion mediated by viral envelope glycoproteins (either baculovirus gp64 or influenza hemagglutinin, HA), includes a low pH triggering event, apparently independent of membrane lipid composition. We have identified a specific stage of fusion which is subsequent to the low pH-induced changes in HA and gp64 conformation, but precedes hemifusion and fusion pore formation. This stage, actual membrane merger, does not require low pH but is still protein-dependent and may be inhibited by trypsin and N-ethylmaleimide (gp64-mediated fusion) and by proteinase K (hemagglutinin-mediated fusion). Membrane merger can be reversibly modulated by changes in membrane lipid composition in a specific correlation with the effective molecular shape of the lipids. Cone-shaped lipids (e.g., oleic acid, OA) promoted, and inverted cone-shaped lipids (e.g., lysophosphatidylcholine, LPC) inhibited fusion, if added to the contacting monolayers of the fusing membranes. These effects are additive, since the exogenous lipids cancel each other if added simultaneously. Addition of LPC caused an increase in the number of fusion proteins that must undergo a low pH conformational change in order to drive fusion. In contrast, OA decreased the number of the activated fusion proteins required. Our data indicate that the different membrane fusion systems including fusion of purely lipid bilayers, involve a consecutive formation of transient and local highly bent intermediates, stalks and pores. Energy of these intermediates and, consequently, the rate and extent of fusion depend on the propensity of the corresponding monolayers of membranes to bend in the required directions.

Project Description:Objectives:

Disparate cell biological reactions converge to the common stage of membrane fusion mediated by the specialized "fusion" proteins, which change their conformation upon interaction with specific triggers. For some examples of the "simplest" class of fusion reactions, used by enveloped viruses to enter host cells, fusion proteins are reliably identified. However little is known about the stage of the biological fusion which is prior to actual membrane merging but subsequent to conformational changes of the fusion protein upon its activation. Since this stage should involve some rearrangements of membrane lipid bilayers, the investigation of the effects of membrane lipids on biological fusion may bring the useful information about the structures of lipid-involving fusion intermediates. In the present work we have used the specific alterations of membrane lipid composition to isolate and characterize the lipid-sensitive stage of cell-cell fusion mediated by different viral glycoproteins. To investigate the lipid rearrangements in fusion we studied the fusion of phospholipid vesicles to planar phospholipid bilayer membranes.

Methods Employed:**Cell-cell fusion-**

1) *Baculovirus gp64-mediated fusion.* Fusion of Sf 9 cells infected by baculovirus was characterized by counting cells in syncytia; by fluorescent microscopy observations of the fluorescent membrane dye redistribution between labeled and unlabeled cells; and by electron microscopy.

The low pH-induced conformational changes of baculovirus fusion protein were confirmed by immunofluorescence microscopy using monoclonal antibodies AcV1 to gp64 generously provided by Dr. G. Blissard (Cornell University, Ithaca). These antibodies specifically bind only to the original, 'neutral' pH form of gp64.

2) *Influenza hemagglutinin-mediated fusion.* HAb2 cells, the line of stably transfected NIH-3T3 fibroblasts expressing the A/Japan/305/57 strain of influenza virus HA were grown to ~80% confluency in DMEM medium supplemented with 10% fetal bovine serum in 5% CO₂, 37°C. CHO-K1 cells expressing the X:31 strain of influenza virus HA (referred to as HA300a cells) or expressing the GPI-anchored X:31 HA (BHA-PI cells) were generously provided by Dr. J. White (University of Virginia, Charlottesville). Cells were grown to ~80% confluency in the medium supplemented with 400 μ M L-methionine sulfoximine. The medium for BHA-PI cells was supplemented with 250 μ M 1-deoxymannojirimycin. The expressed HA0 was cleaved into its fusion-competent HA1-S-S-HA2 form with 5 μ g/ml trypsin in the presence of 250 μ g/ml neuraminidase for 10 min at room temperature.

Human erythrocytes were labeled with fluorescent membrane dye R18 or with fluorescent water-soluble dye, carboxyfluorescein. In some experiments erythrocytes were loaded by lauroyl LPC.

Fusion of phospholipid vesicles to planar lipid bilayer-

Phospholipid vesicles from a mixture of egg phosphatidylcholine with Rhodamine-Phosphatidylethanolamine (R-PE) loaded or not loaded with 200 μ M calcein were prepared daily by a sonication-freeze-thaw technique. The vesicles obtained by this technique have diameters ranging from 20 to several hundred nanometers with the average vesicle diameter estimated to be ~70 nm. Giant unilamellar vesicles (2-10 μ m diameter) loaded with channel-forming antibiotic nystatin (10 μ g/ml), and 200 mM calcein, 5 mM n-propyl gallate and 10 mM MES, pH 6.7 were formed from the mixture of asolectin, ergosterol and R-PE (70:20:10, w/w %) by the technique described in (W.D. Niles, F.S. Cohen, *J. Gen. Physiol.* 90, 703 (1987)).

Solvent-free planar phospholipid bilayers were formed by the Montal-Rudin technique across a ~200 μ m diameter hole in Teflon partition in a special chamber for fluorescent microscopy similar to that described (W.D. Niles, F.S. Cohen, *J. Gen. Physiol.* 90, 703 (1987)). This chamber was placed between an objective and condenser lens mounted on a microbench system. Lipid vesicles were pressure injected toward the planar bilayer so that a number of vesicles became attached to the membrane. Stirring of *cis* compartment removed unbound liposomes. A 10x, 0.21 N.A. lens was used for observation of planar bilayer under bright field illumination and a 40x, 0.5 N.A. lens (Nikon) was used as an objective for fluorescence. Electrical measurements on planar bilayers were performed as described.

Major Findings:

We identified and characterized the lipid-involving intermediate in a viral fusion pathway. This fusion intermediate ('activated state') appear to be common for baculovirus gp64-mediated fusion, for fusion mediated by HA of different strains of influenza (namely Japan and X-31 strains), and for hemifusion mediated by GPI-anchored HA. The intermediate is downstream of the fusion protein transition from initial, neutral pH conformation to the activated one which requires low pH application to contacting membranes and is confirmed by functional, immunological and biochemical approaches. These conformational changes are known to dramatically increase the membrane lipid bilayers capability to fuse. The presence of the activated proteins is required not only to reach the 'activated state' but also to support the membranes commitment to fuse. However, the actual membrane merger and fusion pore formation requires the 'permissive' lipid composition. The lipid effects on this fusion stage correlate with their ability to support the formation of non-bilayer structures of different net curvature. Inverted cone-shaped micelle-forming lysolipids inhibit fusion. Cone-shaped *cis*-unsaturated fatty acids, known to promote inverted hexagonal phase formation, promote viral fusion. The effects of these lipid inhibitors and promoters are additive. Oleoyl LPC and oleic acid compensate the effects of each other upon their incorporation.

Biological fusion involves the distinct stages of the merger of membrane lipid bilayers (including, at least in some systems, the stage of membrane hemifusion, i.e., merger of only outer membrane monolayers), fusion pore formation and expansion. We have found that all these fusion stages can be detected and investigated in fusion of purely lipidic bilayers. Interaction of a giant lipid vesicle with planar lipid bilayer was studied using the combination of fluorescence microscopy and electrophysiological recording to follow the mixing of membranes and their aqueous contents. Hemifusion was observed as membrane dye transfer from liposome to planar bilayer with no release of water-soluble marker of the liposome content. The appearance of small and short-lived fusion pores was detected as conductance flicker reflecting the transient connection between planar bilayer and highly conductive liposome containing nystatin channels. Hemifusion and fusion pore formation and expansion were studied as a function of the membranes lipid composition, divalent cations presence and osmotic tension applied.

Significance to Biomedical Research and the Program of the ICD:

We have identified and isolated a stage of biological fusion which is prior to membranes hemifusion, but subsequent to the triggering event and to the change in fusion protein conformation upon its activation. The dependence of fusion on membrane lipid composition suggest that different fusion systems including fusion of purely lipid bilayers, involve formation of the similar bent intermediates and transient lipidic connections between fusing membranes. Fusion proteins may control the energy of these intermediates by altering the bending propensity of membrane monolayers. The better understanding of the protein-lipid interactions, which underlie membrane merger reaction in a variety of fundamental cell biological processes, may provide the new way to control these processes.

Proposed Course:

1. To study the dependence of the lipid sensitivity of viral fusion on the membrane density of the activated fusion proteins. We are planning to compare the results of these experiments with predictions of the theoretical model assuming that fusion proteins alter the elastic properties of membranes.
2. To compare interaction of two lipid bilayers with interaction of a lipid bilayer with biological membrane.
3. To compare the properties of fusion intermediates arrested by lysolipids with those arrested by temperature decrease.

Publications:

Chanturiya A. A comparative study of planar lipid membranes formed by Montall-Mueller and Mueller-Rudin techniques, *Membr Cell Biol* 1995, in press.

Chernomordik L, Chanturiya A, Green J, Zimmerberg J. The hemifusion intermediate and its conversion to complete fusion: regulation by membrane composition, *Biophys J* 1995, in press.

Chernomordik L, Kozlov M, Zimmerberg J. Lipids in biological membrane fusion, *J Membr Biol* 1995;146:1-14.

Chernomordik L, Leikina E, Cho M, Zimmerberg J. Control of baculovirus gp64-induced syncytia formation by membrane lipid composition, *J Virol* 1995;69:3049-3058.

Chernomordik L, Zimmerberg J. To bend the membranes to the task, *Curr Opin Struct Biol*, 1995, in press.



DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER
Z01 HD 00093-21 OSD

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Mechanism of Action of Nerve Growth Factor

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

PI: G. Guroff, Ph.D. Head, SGF OSD, NICHD

Others: (See Attached)

COOPERATING UNITS (if any)

(See Attached)

LAB/BRANCH

Office of the Scientific Director

SECTION

Section on Growth Factors

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

11.25

PROFESSIONAL:

8.75

OTHER:

2.50

CHECK APPROPRIATE BOX(ES)

- ☐ (a) Human subjects ☐ (b) Human tissues ☒ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unrounded type. Do not exceed the space provided.)

Nerve growth factor (NGF) was the first recognized and is now one of the best characterized of the peptide growth factors. It acts on various elements in the nervous system and on a number of other cell types as well. It is a member of a newly discovered family of homologous growth factors, called the neurotrophins, now five in number, that supports a wide variety of neural cells. There are, in addition, a number of other peptide growth factors that are known to act on the nervous system. NGF is required for the survival and development of sympathetic and sensory neurons. It also is involved in the support of several other cell types, including specific populations of neurons in the central nervous system, the cells of the adrenal medulla, and a number of tumors as well. The action of NGF on these different cells is initiated by its binding to specific receptors. These are made up of two separate proteins, the product of the trk protooncogene and a lower-affinity site known as p75. The binding activates one or more signal transduction pathways that lead to alterations in the phosphorylation and, consequently, the function of key proteins in the cell and to changes in the expression of specific genes. These changes in protein function and in gene expression, caused by the changes in phosphorylation, are the mechanism by which NGF exerts its effects on its target cells. Much of the work leading to this concept has been done with the PC12 pheochromocytoma, a cell line derived from a tumor of the rat adrenal medulla. This clonal line continues to be one of the most informative tools available for the study of NGF and a key model for neuronal differentiation in general. In the presence of NGF, PC12 cells stop dividing, elaborate neurites, become excitable, and will synapse with appropriate muscle cells in culture. Indeed, they change from a rapidly-dividing chromaffin cell to a terminally-differentiated sympathetic neuron within a few days. The changes in phosphorylation that underlie these striking alterations in phenotype occur in virtually every compartment in the cell. Phosphorylation of NGF-stimulated calcium channels appears to regulate the calcium flux across the membrane and, in turn, the intracellular calcium levels, and these levels surely influence the survival of target neurons and may also regulate the ability of the neuron to withstand environmental insults such as occur in stroke. NGF-induced phosphorylation of the elements that control protein synthesis, such as eIF-4E, eEF-2, and S6, almost certainly alters the rate and the specificity of translation. The NGF-induced phosphorylation of specific transcription factors determines which genes are expressed. An understanding of the action of NGF will surely illuminate the control of neuronal differentiation and survival. Clinical interest in these peptide factors is intense because of the role they might play in neurodegenerative diseases.

Others:

G. Dickens	Biological Laboratory Technician	OSD, NICHD
M. Shibutani	Visiting Fellow	OSD, NICHD
T. Tang	Visiting Fellow	OSD, NICHD
W. Zheng	Visiting Fellow	OSD, NICHD
Y. Katagiri (0.75 yr)	Visiting Fellow	OSD, NICHD
M. Whalin	IRTA	OSD, NICHD
M. Monshipouri	IRTA	OSD, NICHD
H. Jiang	IRTA	OSD, NICHD
E. Kim	Pre-IRTA	OSD, NICHD
V. Movsesyan	Exchange Scientist	OSD, NICHD
M. Lavarreda (0.5 yr)	Howard Hughes (High School) Intern	OSD, NICHD

Cooperating Units:

Department of Pathology, Washington University School of Medicine
(J. Milbrandt)

Department of Microbiology and Molecular Genetics, Harvard Medical School
(M. Greenberg)

School of Pharmacy, The Hebrew University of Jerusalem, Israel
(P. Lazarovici)

Department of Internal Medicine, Fukui University, Japan
(T. Mutoh)

Biomimetic Control Research Center, Nagoya, Japan
(Y. Hirata)

National Institute of Neurological Disorders and Stroke
(H. Chin)

Project Description

Objectives:

The purpose of this project is to provide an understanding of the biochemical, molecular, and cellular actions of nerve growth factor (NGF). Such an understanding is important for a number of reasons. First, the action of nerve growth factor causes the survival and differentiation of many of its targets and information on how it works may shed light on the mechanisms of normal and abnormal neuronal development and differentiation and, as well, on the outlines of cellular differentiation in general. Second, nerve growth factor is the lead member of a newly described class of homologous factors called neurotrophins and is a prototype for any number of related growth factors acting on neural tissue, and an understanding of the actions of nerve growth factor may point the way to an understanding of the actions of these other factors. Third, the actions of nerve growth factor in promoting the survival of certain of its target neurons may provide information of clinical value in cases where neurons die prematurely, such as in Alzheimer's Disease and Amyotrophic Lateral Sclerosis, or where neurons are damaged by their environment, such as in cases of stroke or peripheral neuropathies. Finally, since the actions of nerve growth factor on many of its targets are to decrease cell division, an understanding of the mechanism by which it acts may provide clues as to how the alternate cellular fates, differentiation and division, are controlled, not only in normal cells, but in the several tumor cells on which nerve growth factor acts. Interest in this field has increased dramatically over the last few years, due to a number of observations. First, the understanding that nerve growth factor had actions in the central nervous system indicated that its role was not limited to specific portions of the peripheral nervous system. Then, the discovery of other neurotrophins broadened the field to include virtually all areas of the nervous system and to increasing aspects of normal development as well. Finally, the realization that these peptides could have clinical significance, in both neurological diseases and brain injury, has attracted significant commercial activity.

The approach taken in this laboratory has a number of separate, but related foci. First, we are interested in the reactions immediately following the combination of nerve growth factor with its receptors. These experiments involve an identification of the second messengers generated by such a combination and an exploration of the biochemical or biophysical reactions that control their levels. Second, we are concerned with the effect of these receptor-level events on protein phosphorylation reactions in the cell. Here we attempt to construct and dissect cell-free systems that reflect the actions of nerve growth factor on the intact cell. Third, we are involved in studies that explore the changes in the expression of specific genes responsible for the nerve growth factor-induced alterations in the properties of the cells, and how these gene-level changes are regulated. In these experiments we make use of molecular biological manipulations to alter the action of the trans-acting factors that control the expression of specific genes and to determine the effect of these alterations on the actions of nerve growth factor. Finally, we are attempting to understand how these biochemical and molecular changes alter the phenotype of the cell. In these studies we try to determine such things as the role of nerve growth factor-induced calcium uptake in the survival of the cell or the role of nerve growth factor-induced decrease in mitogen receptors in the control of cell division and neuronal maturation. The overall aim is to understand how nerve growth factor acts in biochemical, molecular, cellular, and clinical terms.

Methods Employed:

The methods used include the standard techniques of cell culture and analysis, almost exclusively with the nerve growth factor-responsive cell line PC12 and its variants. Cell fractionation and enzyme assay are done routinely. Enzyme purification relies heavily on the use of fast protein liquid chromatography (FPLC). Protein phosphorylations are analyzed by polyacrylamide gel electrophoresis, either in SDS or acid-urea buffer systems, and phosphopeptide mapping is carried out by two-dimensional analysis or by HPLC. Peptide sequencing has permitted the identification of proteins, the phosphorylation of which is altered in nerve growth factor-treated cells. Polyclonal antibodies are prepared against proteins purified in the laboratory or against peptide sequences of interest prepared synthetically. These antibodies are also used to detect, by western blotting, or to isolate, by immunoprecipitation, proteins of interest. Commercial antibodies are frequently used immunohistochemically to localize specific proteins in or on cells. Various proteins are radioiodinated in order to measure the binding of these ligands to their specific receptors on the cell surface. Antibodies or specific toxins are introduced into cells using electroporation, or by permeabilization of the cells with mild detergents. Radioactive ligands have been used

to evaluate the role of specific calcium channels in the actions of nerve growth factor. Various toxins are evaluated for their effects on nerve growth factor-activated ion channels. Fluorescent intracellular calcium-binding dyes have been most valuable in the evaluation, by fluorescence-activated cell analysis, of the effects of growth factors on channel function. Various combinations of transfection or infection are employed to introduce specific genetic elements into the cells in order to evaluate their roles in the differentiation produced by nerve growth factor. The use of antisense probes has allowed the deletion of selected proteins and the construction of dominant-negative transfectants has permitted an analysis of the role of certain signal transduction molecules in nerve growth factor action. Site-directed mutagenesis is now employed to alter specific residues in the nerve growth factor receptors in order to determine the roles of these residues in signaling. Nucleic acid hybridizations have been of great value in studying, on northern blots, the patterns of gene expression following treatment of the cells with different growth and differentiation regulators. Most recently, translation of specific messages using a fortified rabbit reticulocyte system or an S10 extract of PC12 cells has become a high priority effort in the group.

Major findings:

Our effort to elucidate the molecular events mediating the actions of nerve growth factor on its target cells has led to the following specific advances during the past year.

- 1) Arachidonic acid is one of the many second messengers whose levels are altered by the combination of nerve growth factor with its receptor on PC12 cells. The biochemical pathway by which this increased arachidonic acid is produced has not been elucidated. Recent evidence implicates protein kinase C in this reaction pathway. Down-regulation of kinase C by long-term treatment of the PC12 cells with the tumor promoter PMA prevents the action of nerve growth factor on arachidonic acid release and inhibitors of kinase C also block this effect. Multiple isoforms of kinase C have been detected in the cells with isoform-specific antibodies and one isoform, protein kinase C alpha, has been identified, through the use of isoform-specific inhibitors and activators, as the one specifically involved in nerve growth factor-stimulated arachidonic acid release. This work highlights the differential effects of the several kinase C isoforms.
- 2) Calcium uptake into PC12 cells is increased by treatment of the cells with nerve growth factor. Furthermore, this action of nerve growth factor depends upon a unique calcium channel and the stimulating action of nerve growth factor on this channel is probably due to a phosphorylation of one or more of the channel subunits. It has been suggested that, since the survival of neurons requires adequate intracellular calcium levels, this action may be one of the mechanisms by which nerve growth factor keeps neurons alive. Present work focusses on the pathway by which this channel is activated and recent data implicate protein kinase C. Long-term down-regulation of kinase C with the tumor promoter PMA prevents the action of nerve growth factor on calcium uptake. Studies currently underway are directed toward the identification of the specific isoform of the enzyme involved and at the determination of the channel specificity of the action of that isoform. Another line of study concerns the possible action of nerve growth factor in the phosphorylation of calcium channel subunits in PC12 cells. Antibodies directed against the beta subunit of the brain-specific L-type calcium channel have co-immunoprecipitated a 61 kDa protein from PC12 cells that is phosphorylated upon treatment of the cells with nerve growth factor. N-terminal amino acid sequence analysis shows that the protein is unique. Appropriate nucleotide probes have been used to extend the sequence and cloning is underway. These experiments may allow a determination of the manner in which a unique nerve growth factor receptor-associated calcium channel is activated and an elucidation of its molecular nature.
- 3) A continuing interest in the nerve growth factor-stimulated phosphorylation of the ribosomal protein S6 has led, in previous work, to the identification of a unique 47 kDa nerve growth factor-stimulated S6 kinase in the cytoplasm of PC12 cells. It now appears that this new kinase is related immunologically to the p70^{S6K} present in many cells. Current work has ruled out the possibility that it is a proteolytic fragment and has focussed attention on the splice variants. The enzyme has proved to be stable and clearly not any of the known PC12 cell kinases. Purification of the enzyme is proceeding. In another study, it has been shown that there is a nerve growth factor-stimulated S6 kinase in the nuclei of PC12 cells. This kinase has been identified as the nuclear form of p70^{S6K}, p85^{S6K}, and it has been found that it is inhibited by both rapamycin and wortmannin. The presence of S6, itself, in the nucleus of PC12 cells has been observed and experiments are underway to try to understand the role of this protein in nuclear function. The study of the S6 kinases in PC12 cells has important implications for the action of nerve

growth factor on protein synthesis and, in turn, on the phenotypic changes nerve growth factor induces in its target cells.

- 4) The finding, some years ago, that PC12 cells respond both to the differentiating agent nerve growth factor and the mitogen epidermal growth factor, has led to important studies on comparative signal transduction and to equally significant experiments on receptor interplay. The latter are based on the finding that differentiation with nerve growth factor causes the down-regulation of the epidermal growth factor receptor. Recent data have shown that mRNA levels for this receptor are unchanged even though receptor levels are markedly decreased. For this and other reasons, present work focusses on a possible nerve growth factor-induced decrease in the translation of the epidermal growth factor receptor mRNA. Two different translation systems have been used, one a fortified rabbit reticulocyte lysate and the other an S10 system from the PC12 cells themselves. Using these systems it has been possible to translate epidermal growth factor receptor mRNA and the effort now is to obtain clones with sufficient flanking information to determine the mode by which this translation is regulated. A parallel line of study involves the translation of the mRNA for erbB2, a receptor that is homologous with the epidermal growth factor receptor, one which is also down-regulated by treatment of the cells with nerve growth factor, and one for which mRNA levels are also unchanged during the differentiation. Yet another series of experiments has shown that erbB3, which has just been identified on these cells, is also down-regulated during differentiation. These studies give promise of elucidating the molecular mechanism by which the levels of these related, clinically-relevant receptors are regulated.
- 5) Although it is known that nerve growth factor alters the cell cycle of PC12 cells and stops them from dividing, the exact alterations that nerve growth factor produces are not clear. Present work is pointed at two questions. First, are the cells equally sensitive to nerve growth factor at all stages of the cell cycle and, second, what changes does nerve growth factor produce in the cycle. Previous work on the first question has pointed to the G1/S border as the most sensitive point for such responses as the induction of ornithine decarboxylase and the increased expression of c-fos. Present studies are focussed on some of the longer-term events initiated by nerve growth factor. Information on the second point has been obtained within the last several months. It has been shown, using flow cytometry, that nerve growth factor treatment causes an increase in the number of cells in G2 and a decrease in the number in G1. This alteration is accompanied by changes in the levels of cyclins consistent with the alterations in the cell population, i.e., decreases in both cyclins B and F. Other studies have indicated that the application of mitogens such as insulin-like growth factor I to cells fully differentiated with nerve growth factor is able to produce increases in DNA synthesis without significant changes in cell number. These studies are designed to reveal the fundamental alterations in the cell cycle that lead to nerve growth factor-induced differentiation of these cells.

Overall, research conducted under this project has shown that the nerve growth factor-induced phosphorylation of key proteins in nerve growth factor-sensitive cells is a major part of the mechanism by which nerve growth factor acts. These changes in phosphorylation and the consequent changes in function are the sum and substance of the phenotypic changes that nerve growth factor produces. Our recent studies have begun to identify specific gene products, e.g., the nerve growth factor-sensitive calcium channel, protein kinase C, the ribosomal protein S6, the epidermal growth factor receptor, that play crucial roles in these alterations in phenotype. Further, some of the newer experiments are designed to explore how these specific gene products contribute to the global changes in phenotype the cells experience under the influence of nerve growth factor.

Significance to Biomedical Research and the Program of the Institute:

The normal development of the nervous system is both intricate and precise and is integrated in both time and space. The appropriate neural connections must be made in exactly the right place and at exactly the right developmental times. Neurons reach their appropriate locations, stop dividing, and differentiate on an exquisitely complex and exact timetable and redundancies are eliminated by programmed cell death. It has become clear that the neurotrophins and other neural growth factors are key players in the control of these processes. Through chemotactic mechanisms they encourage the correct connections; through differentiative processes they inhibit cell division; through competition for their limited quantities they elicit programmed cell death. An understanding of the detailed mechanisms of these interactions will elucidate a number of fundamental biological processes. Such an understanding, also, will permit insights into what happens when these processes misfire. These insights should

allow the formulation of clinical approaches toward the correction of these aberrant states. Through an understanding of such fundamental biological phenomena as process formation, chemotactic movements, and programmed cell death it should be possible to consider solutions to complex problems such as mental retardation. By exploring the control of normal neural cell death during development, information should be developed that is relevant to conditions involving abnormal cell death, such as Alzheimer's Disease, peripheral neuropathies, and stroke. The clinical use of growth factors in such conditions is a present reality and, since the rate of discovery of new growth factors is on an ascending curve, it is also an abundant promise for the future.

Proposed Course:

We will continue to explore the events immediately following the binding of nerve growth factor to its receptor, the pathways by which the signal generated at the receptor is transmitted throughout the cell, the changes these signals produce in the nucleus, and the role of specific gene products in the overall phenotypic changes that nerve growth factor produces. We will try to (1) identify the specific isoform of protein kinase C that is involved in the nerve growth factor-induced stimulation of calcium uptake; (2) clone the unique 61 kDa protein that is co-immunoprecipitated with the beta subunit of the brain-specific L-type calcium channel and that is phosphorylated upon treatment of PC12 cells with nerve growth factor; (3) purify and characterize the unique nerve growth factor-activated 47 kDa S6 kinase from PC12 cells and determine the sites at which it phosphorylates S6; (4) translate clones of the epidermal growth factor receptor, erbB2, and erbB3 that contain both 5' and 3' flanking information in the S10 fractions from control PC12 cells and from PC12 cells differentiated with nerve growth factor; (5) perform site-directed mutagenesis on the transmembrane regions of the nerve growth factor receptors and determine the effect of these mutations on the ability of nerve growth factor to act on the cells; (6) overexpress erbB2 in PC12 cells and determine if nerve growth factor down-regulates this overexpressed receptor; (7) transfect PC12 cells with the regulatory region of the kinase B-raf and determine the effect of this dominant-negative element on signal transduction and gene expression; (8) transfect PC12 cells with the DNA-binding region of the transcription factor NGFI-B and explore the effect of phosphorylation of this domain on its transcriptional activity; (9) employ the down-regulation of the low-affinity nerve growth factor receptor with dexamethasone to determine if this receptor is involved in cell survival; (10) elucidate the effect of nerve growth factor on the cell cycle of PC12 and correlate the changes in cell population with biochemical alterations in cyclin levels.

Publications:

Hirata Y, Whalin M, Ginty DD, Xing J, Greenberg ME, Milbrandt J, Guroff G. Induction of a nerve growth factor-sensitive kinase that phosphorylates the DNA-binding domain of the orphan nuclear receptor NGFI-B. *J Neurochem* (in press).

Nikodijevic B, Aschkenasy M, Dickens G, Lachance C, Guroff G. Characteristics of the K-252a-induced increase in calcium uptake in PC12 cells. *J Neurosci Res* 1995;40:494-8.

Oshima M, Weiss L, Dougall WC, Greene MI, Guroff G. Down-regulation of c-neu receptors by nerve growth factor in PC12 cells. *J Neurochem* 1995;65:427-33.

Oshima M, Hirata Y, Guroff G. The nerve growth factor family. In: LeRoith D, Bondy C, eds. *Growth factors and cytokines in health and disease*. Greenwich, CT, Jai Press, Inc (in press).

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 01500-13 OSD

PERIOD COVERED

October 1, 1994 through September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Studies on DNA Replication, Repair, & Mutagenesis in Eukaryotic & Prokaryotic Cells

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

P.I.: A.S. Levine, M.D., Head, SDRRM, OSD, NICHD

E. Chumakov, Ph.D. - IRTA Fellow	M. McLenigan	- Chemist
D. Ennis, Ph.D. - Sr. Staff Fellow	M. Protic, Ph.D.	- Sr. Staff Fellow
J. Hauser	- Microbiologist	V. Rasic-Otrin, Ph.D. - IRTA Fellow
G. Kopsidas, Ph.D. - Exchange Scientist	O. Studitskaia, Ph.D.	- Visiting Fellow
J. McDonald, Ph.D. - IRTA Fellow	R. Woodgate, Ph.D.	- Sr. Staff Fellow

COOPERATING UNITS (if any)

NLM (E. Koonin & J. Wootton); NIDDK (E. Ackerman); FDA (W. Koch); Columbia University, New York (T. Peat & W. Hendrickson); MRC, England (R. Wood)

LAB/BRANCH

Office of the Scientific Director

SECTION

Section on DNA Replication, Repair, and Mutagenesis

INSTITUTE AND LOCATION

NICHD, NIH, Bethesda, Maryland 20892

TOTAL STAFF YEARS:

11.0

PROFESSIONAL:

9.0

OTHER:

2.0

CHECK APPROPRIATE BOX(ES)

☐ (a) Human subjects ☒ (b) Human tissues ☐ (c) Neither
☐ (a1) Minors
☐ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

Studies on the mechanism of mutagenesis in prokaryotic cells have focused on the roles of the RecA and UmuDC-like mutagenesis proteins (which facilitate DNA polymerase bypass of unrepaired DNA lesions). We have overproduced, purified and crystallized the *E.coli* UmuD' protein. The structure was refined to 2.5 Å and analyses revealed that in addition to forming a molecular dimer with itself, the amino terminal of UmuD' can interact with the amino terminal of another "molecular dimer" to form an extended polymer ("filament") structure. While deletion of the N-terminal of UmuD' still allows the protein to form the "molecular dimer", it precludes the formation of the polymer. The N-terminal deletion mutant of UmuD' demonstrates a greatly reduced ability to bind to the RecA nucleoprotein filament and explains why cells expressing the mutant (together with UmuC) are non-mutable. In other studies, construction of chimeric UmuC proteins were generated to investigate the poorly mutable phenotype associated with *S.typhimurium*. These constructs revealed that alterations in the region located between *S.typhimurium* UmuC residues 26-59 are most likely the cause of the poorly mutable phenotype. We have also investigated the *in vivo* stability of the Umu proteins in *E. coli*. UmuC appears to be inherently unstable, but is partially stabilized in the presence of UmuD. The mutagenically active UmuD' appears to stabilize UmuC further; moreover, UmuC was stabilized still further when RecA protein was constitutively expressed in its activated state. Studies with *Xenopus laevis* demonstrated that while oocytes can efficiently replicate undamaged single-stranded DNA, they are unable to replicate DNA containing adducts. This replication arrest was alleviated in progesterone-matured oocytes and in oocytes microinjected with mRNAs encoding the prokaryotic UmuD'C or MucA'B mutagenesis proteins, suggesting that the basic mechanisms of mutagenesis are highly conserved between prokaryotic and eukaryotic cells.

In a second project, we continued studies on a protein (UV-DDB) which binds to 6/4 photoproducts in UV-damaged human DNA. Purified UV-DDB was not essential in an *in vitro* nucleotide excision repair system, but it may be required *in vivo*. UV-DDB moves to a tight association with damaged DNA upon UV treatment of cells; RPA, which also redistributes after UV, is also present in the UV-DDB/DNA complex. The interaction of DDB and RPA enhances the DNA binding of either alone. This interaction, similar to XPA + RPA, places UV-DDB in the initial damage recognition step of DNA repair.

PROJECT DESCRIPTION

OBJECTIVES:

A. PROKARYOTIC MUTAGENESIS

Under optimal conditions the fidelity of DNA replication in *Escherichia coli* is extremely high, on average, only one error occurs in every 10^{10} bases replicated. However, upon exposure to a variety of DNA damaging agents, such as ultraviolet light, many *E. coli* strains exhibit a dramatic increase in the mutation frequency. Genetic and physiological experiments have indicated that DNA polymerase III holoenzyme (the main replicative enzyme of *E. coli*), RecA, and the UmuDC-like mutagenesis proteins are directly required for this process; all of which are induced as part of the cell's RecA-LexA controlled multi-gene SOS response to DNA damage. During the mutagenic process, RecA mediates the posttranslational cleavage of the UmuD-like proteins, leading to their activation for mutagenesis functions. The actual mechanism by which the UmuD'C-like and RecA proteins promote mutagenesis remains to be elucidated, although it is believed that they act together to promote translesion synthesis by DNA polymerase III holoenzyme. Thus, our continuing biochemical and genetic studies of the UmuDC-like proteins will provide us with a better understanding of the mechanisms of mutagenesis in *E.coli*, and in all likelihood, higher organisms.

B. DNA REPAIR

DNA repair is an essential cellular process which preserves genomic integrity and programmed gene expression, and protects a cell from neoplastic transformation and death. Various DNA lesions, induced by endogenous or environmental agents, are efficiently recognized and repaired by different repair pathways in both prokaryotic and eukaryotic cells. The importance of the repair processes for human health has become apparent from studies on human hereditary diseases associated with a deficiency in DNA repair: e.g., a high susceptibility of xeroderma pigmentosum (XP) patients to sun-induced skin cancer is strongly correlated with the hypersensitivity of XP cells to UV light, and their decreased capacity to excise UV-induced photoproducts. Recently, several genes and gene products that are defective in XP cells have been isolated and cloned from normal mammalian cells. This small subset of about 25 gene products implicated in eukaryotic nucleotide excision repair (NER) has uncovered the complexity of the DNA repair apparatus but provided only a partial understanding of mammalian NER processes. In our studies, we have focused on the early step of NER — lesion recognition — with the goal of identifying proteins that interact with damaged DNA and which may form a repair-initiation complex:

- 1) The absence (or altered regulation) of a damaged DNA-binding (DDB) complex in DNA repair-deficient cells is used as an indicator of the involvement of this binding activity in NER. DDB complexes are being purified and biochemically characterized, and their corresponding cDNAs isolated and expressed in various backgrounds to define more precisely their role in NER and their mechanism of regulation.

- 2) Homologs of novel mammalian DDB proteins are being identified and cloned from lower eukaryotes to determine their evolutionary conservation and preservation of DNA-binding structure, and to better address their putative role in other cellular processes.
- 3) The effect of cellular or exogenous factors on the recognition of DNA lesions as well as the capacity for NER is being examined to expand our understanding of the mechanisms and factors that modulate these processes.
- 4) To establish a simple DNA-binding assay for the diagnosis of repair-deficient individuals, we are testing the hypothesis that the recovery of DDB activity after exposure of cells to DNA-damaging agent(s) positively correlates with the cells' capacity for NER.

METHODS EMPLOYED

A. *PROKARYOTIC MUTAGENESIS*

Standard methods of molecular biology, biochemistry and molecular genetics are employed in this work. Recombinant DNA clones have been constructed for the overproduction and characterization of the UmuD'-like and UmuC-like mutagenesis proteins. Plasmid DNAs are typically isolated using Qiagen columns and analyzed further by restriction enzyme digestion. Recombinant clones of interest are amplified by transformation of competent bacteria. Overproduced proteins are typically purified using ion-exchange, gel-filtration and affinity chromatography. Proteins are separated and analyzed using SDS-PAGE. Immuno-detection of proteins is achieved using the Western blotting technique, with visualization of proteins using a chemiluminescent substrate. Protein-protein interactions are studied directly using affinity chromatography and chemical cross-linking, or indirectly using the yeast "two-hybrid" system. Protein-DNA interactions are studied using an agarose based DNA mobility shift assay as well as filter binding assays. Novel bacterial strains used in this study are constructed using generalized phage P1 transduction.

B. *DNA REPAIR*

Standard methods of cell culture and cell fractionation are used to prepare whole-cell, nuclear and cytoplasmic extracts. Liquid and affinity chromatography, and SDS-PAGE are used to purify DDB proteins. A damaged-DNA band-shift assay is employed to detect and characterize DDB proteins. DNA-binding substrates, specific DNA-cleavage reactions and protein fractions are analyzed by PAGE methods. PCR technology is applied to isolate cDNAs encoding DDB proteins. Standard methods of molecular biology are used to clone and sequence cDNAs. Cell transfection techniques are utilized for *in vivo* DNA repair studies and selection of cells carrying cloned DDB genes. Immuno-detection by Western blotting, and *in situ* by EM-immunochemistry, is employed to identify and localize DDB proteins, to quantitate DDB

proteins in various subcellular fractions, and to identify proteins that interact with DDB. Enzyme and protein assays are carried out by standard biochemical techniques.

MAJOR FINDINGS

A. PROKARYOTIC MUTAGENESIS

Crystallographic studies of the *E. coli* UmuD' protein

In an attempt to increase our understanding of the mutagenic process, we sought to determine the molecular structure of the mutagenically active UmuD' protein. While we obtained crystals of UmuD' and gathered a native data set, we were unable to obtain a 3-dimensional structure by conventional methods of isomorphous replacement. As an alternative approach, we purified UmuD' from bacteria cells grown in a defined medium containing Seleno-methionine. (Substitution of Selenium for Sulfur in the methionine amino acid residue provides sufficient phase discrimination to obtain the 3D structure of seleno-methionine substituted proteins). Unfortunately, we were unable to obtain crystals of the seleno-methionine substituted wild-type UmuD' protein. UmuD' normally contains three methionine residues and one of these is the penultimate residue of the protein. One plausible hypothesis to explain the lack of seleno-methionine substituted UmuD' crystals is that this residue is exposed on the surface of the protein so that it becomes oxidized and consequently precludes crystallization of the protein. To overcome these possibilities, we used site directed mutagenesis to change the penultimate methionine residue to those of threonine and valine (residues found in homologous UmuD'-like proteins). Both of these UmuD' mutants were fully functional for mutagenesis and in fact were 2-3 fold better at promoting mutagenesis than the wild-type protein. Both of the seleno-methionine substituted UmuD' mutant proteins crystallized but the UmuD'-Thr mutant gave larger and better structure crystals and was therefore chosen for further studies.

The protein was crystallized by the hanging drop method with a final protein concentration of 12 to 15 mg/ml in: 600 mM LiSO₄, 20 mM MgCl₂, 100 mM cacodylate buffer pH 5.8, 5 mM DTT at 20°C. The protein crystallizes in space group P4₁2₁2 with cell dimensions a=b=52.8 Å, c=160.1 Å. The protein adopts an apparently novel fold in that it is a tight beta structure with an extended amino terminus that has a single alpha helix encompassing residues 39 through 45. One turn, residues 61 through 64, adopts a ₃₁₀ helix conformation, and the rest of the structure is best described by seven beta strands connected by short loops. The seven strands are oriented to form three beta sheets; one with four strands and two with three strands. These sheets form the main body of the protein, which has two separate hydrophobic cores. The last beta strand, B7, is quite long and pairs in two different beta sheet formations. This strand is also interesting in that it slides between two other beta strands in the structure as described below.

The crystal is formed by two different kinds of protein interactions, one we call the molecular dimer and the other we call the extended or filament dimer. The asymmetric unit of the crystal is the molecular dimer. This dimer interface is primarily hydrophobic and consists in part of residues val⁵⁴, ile⁸⁷, phe⁹⁴ and phe¹²⁸, with the two phe⁹⁴ residues, one from each monomer,

being stacked on top of each other. There is also a salt bridge between the conserved residues lys⁵⁵ and glu⁹³ on both sides of the interface. To test the hypothesis that the molecular dimer does indeed represent the true dimer structure in solution, we used standard PCR technology to generate a mutant of UmuD' that lacked the amino terminal residues 25-45. In our model, this N-terminal deletion mutant is not expected to form the filament dimer, but would be expected to still form the molecular dimer. This N-terminal UmuD deletion mutant exhibited purification characteristics similar to the wild type protein suggesting that it is correctly folded and both gel filtration and native acrylamide gel electrophoresis suggest that it is indeed a dimer in solution. We have examined this further by physically cross-linking the subunits with glutaraldehyde and then analyzing the products on denaturing SDS-gels. As expected, like the wild-type UmuD⁺ and UmuD' proteins, the N-terminal deletion mutant formed homodimers. While wild type UmuD' forms heterodimers with UmuD⁺, the N-terminal deletion mutant failed to form these dimers. We are presently trying to determine the basis of this phenotype.

The filament dimer is formed by interactions at the amino and carboxyl termini of the protein. Residues 134 through 136 of the carboxyl terminus form hydrogen bonds across the interface with the related residues 134' through 136'. This interaction extends two three stranded beta sheets into a longer six stranded beta sheet. A mutation at residue 138, methionine to either threonine or valine, is a gain of function mutant, showing higher levels of mutagenesis than the wild type protein. Although this residue is not clearly seen in the electron density maps, it is certainly in this interface. It seems likely that these mutants are forming better contacts either between the two UmuD' molecules or between the UmuD' and RecA proteins, thus forming a stronger filament.

At the amino termini, the two alpha helices cross over each other with the side chain oxygen of asn⁴¹ making a hydrogen bond with the main chain nitrogen of leu⁴⁰ (2.7 Å apart). There are also several hydrophobic contacts in this area with leu⁴⁰ being close to leu⁴³, leu⁴⁴, ile⁴⁵ and ile⁷³ on the other chain. Leu⁴⁴ is also close to val¹³⁵ and leu⁴³ interacts with phe⁵³ and leu⁷¹. The extensive interactions in this interface suggest that this is more than just a crystallographic artifact. The biological importance of this extended structure maybe explained in light of Gimble and Sauer's observations that certain mutations in the λcI repressor (corresponding the amino terminus of UmuD) alter the ability of λcI to interact with RecA. To test the hypothesis that the extended N-terminal may contact RecA, we analyzed the ability of the wild-type and the N-terminal deletion mutant to interact with a RecA nucleoprotein filament. While wild-type UmuD' clearly has some affinity for the RecA filament, the N-terminal deletion mutant which lacks residues 25-45 of UmuD clearly showed a reduced affinity for the RecA-filament, suggesting that residues 25-45 are indeed critical for a UmuD'-RecA-DNA interaction. As one might expect, in the absence of the RecA-DNA interaction, the N-terminal deletion mutant protein fails to be targeted to DNA and, as a consequence, cells expressing this protein together with UmuC are rendered phenotypically non-mutable.

Translesion DNA synthesis on *Xenopus laevis* oocytes

Closed circular ss DNA is efficiently replicated to ds DNA in *Xenopus* oocytes, eggs and nuclear extracts. Oocytes do not replicate injected ds DNA, whereas *Xenopus* eggs and egg

extracts replicate ~5% of injected ds DNA. Since oocytes and nuclear extracts are also both extraordinarily efficient at ds DNA repair, we investigated how oocytes and nuclear extracts might replicate/repair UV damaged ss DNA. ss DNA was irradiated with UV at various dosages injected into oocytes with a ^{32}P -labeled dCTP and analyzed. Virtually no DNA synthesis was observed on UV-irradiated circular ss DNA containing as few as ~2-3 cyclobutane dimers (CPD)/ss M13, while unirradiated ss DNA replicated to full length forms I and II. No partial replication products were observed, even when analyzed on 5-20% denaturing PAGE gels; nor does replication occur on UV-irradiated linear ss M13. Furthermore, photoreversal of UV lesions did not occur because such repaired DNA should have then replicated to forms I and II. Southern blot hybridizations made from both TBE and alkaline gels demonstrate that the irradiated circular ss DNA is not degraded by the oocyte. To further characterize the replication arrest mechanism, we next determined if irradiated ss M13 prevents replication of unirradiated ssDNA, i.e. can the arrest inhibit all cellular DNA synthesis? Combining equal amounts of UV-irradiated ϕX174 with unirradiated M13 or vice-versa allows complete replication of the unirradiated DNA. Combining irradiated with unirradiated ss M13 results in replication of the unirradiated portion of M13, so sequence differences between ϕX174 and M13 are irrelevant. Therefore, the oocyte replication arrest mechanism only applies to damaged template molecules. We also examined replication arrest in other *Xenopus* cells. Oocytes are arrested at first meiotic prophase; matured oocytes are released from this arrest, lack a nuclear membrane, and develop as far as metaphase of the second meiotic division. The same batch of oocytes unable to replicate irradiated ss M13 was able to replicate this substrate when it was injected after progesterone-induced maturation. Therefore, the replication arrest exhibited by oocytes with irradiated ssM13 is abrogated during meiosis. These results suggest the emergence in matured oocytes of a replicative bypass pathway. Alternatively, progesterone-induced maturation may effect a displacement of putative anti-replication protein(s).

Since the UmuD'C and MucA'B mutagenesis proteins are thought to act by facilitating translesion DNA synthesis in prokaryotic cells we wanted to determine if they would perform similar roles in a heterologous system. Since producing functional UmuC or MucB has been intractable, the oocyte's translational proficiency was utilized by injecting mRNA transcripts encoding the prokaryotic proteins. Initially, the mRNAs were injected cytoplasmically, and after healing, the oocyte nuclei were injected with label and ss M13 DNA. Combinations of either *mucA'* + *mucB* or *umuD'* + *umuC* facilitated replication of the irradiated ss M13 template. mRNAs injected into the nucleus simultaneously with irradiated ss M13 also facilitated replication, consistent with reports that mRNAs injected into the nucleus are exported quickly to the cytoplasm. Only cognate mRNA combinations promoted significant replication. No replication occurred when non-cognate mRNA partners *mucA'* + *umuD'*, or *umuC* + *mucB*, were injected and only limited synthesis was seen with *umuD'* + *mucB*.

How might UmuD'C and MucA'B proteins alleviate the replication arrest on damaged ss M13 DNA in oocytes? The arrest only functions in *cis*, suggesting that perhaps the bacterial proteins directly interact with the *Xenopus* replication machinery and/or putative *cis* anti-replication protein(s). The form II products synthesized from UV-irradiated ss M13

facilitated by UmuD'C or MucA'B proteins are more diffuse than those made from an undamaged template, suggesting that prokaryotic proteins might not allow extensive synthesis past CPDs in the oocyte. Alkaline agarose gel analysis of replication products on the damaged template demonstrates that the prokaryotic proteins promoted initiation and elongation of DNA synthesis, but a significant portion of the replication products were not full length compared to products from the unirradiated control template. UmuD'C and MucA'B therefore abrogate the replication arrest in oocytes and facilitate some complete bypass synthesis. We have initiated experiments to determine whether the replication observed is translesion or mutagenic.

Characterization of R-plasmid encoded *umu*-complementing genes

In addition to conferring resistances to antibiotics and heavy metals, certain R-factors carry genes involved in mutagenic DNA repair. These plasmid encoded genes are structurally and functionally related to the chromosomally encoded *umuDC* genes of *Escherichia coli* and *Salmonella typhimurium*. Three such plasmid operons, *mucAB*, *impCAB*, and *samAB* have been characterized at the molecular level. Recently, we have identified three additional *umu*-complementing operons from the IncJ plasmid R391, and two IncL/M plasmids, R446b and R471a. The nucleotide sequence of the minimal R-plasmid *umu*-complementing (*rum*) region revealed an operon of two genes, *rumA*_(R391) and *rumB*_(R391), with an upstream regulatory signal strongly resembling LexA binding sites. Phylogenetic analysis revealed that the RumAB_(R391) proteins are approximately equally diverged in sequence from the chromosomal UmuDC proteins and the other plasmid encoded Umu-like proteins and represent a new subfamily. Genetic characterization of the *rumAB*_(R391) operon revealed that in *recA*⁺ and *recA1730* backgrounds, the *rumAB*_(R391) operon was phenotypically indistinguishable from *mucAB*. In contrast however, the *rumAB*_(R391) operon gave levels of mutagenesis that were intermediate between *mucAB* and *umuDC* in a *recA430* strain. The latter phenotype was shown to correlate with the reduced posttranslational processing of the RumA_(R391) protein to its mutagenically active form, RumA'_(R391). Thus, the *rumAB*_(R391) operon appears to possess characteristics that are reminiscent of both chromosomal and plasmid encoded *umu*-like operons.

Molecular analysis of the cloned *umu*-complementing genes from R471a and R446b, two related incL/M plasmids, revealed that they were structurally and functionally more similar to the MucAB proteins from the incN plasmid, pKM101 than to other members of the previously identified Umu-like family. Interestingly, while the *mucAB*-like coding regions of R471a and R446b were highly homologous, the flanking regions were polymorphic. Analysis of these flanking regions revealed an insert in R471a that appears to be a novel retroelement encoding a putative reverse transcriptase (RT). This RT is related to the reverse transcriptases encoded by group II introns but is embedded in a retron-like context. R471a also encodes a putative transposase and sequences that resemble insertion sequences immediately downstream of the *mucAB*_(R471a) operon. Both the RT and transposase elements were absent in R446b. These observations suggest that the *muc*-like genes from R471a and R446b are located within regions of the R-plasmids that perhaps were once (or still are) mobile genetic elements and possibly explains the distribution of *umu*-like genes on R-plasmids and bacterial chromosomes.

Analysis of chimeric *Salmonella typhimurium* UmuC proteins

Unlike *Escherichia coli*, the closely related bacterium *Salmonella typhimurium*, is much less responsive to the mutagenic effects of DNA damaging agents. Previous experiments reported by Sedgwick *et al.*, (Mol. Gen. Genet. (1991) 229:428-436), have suggested that these phenotypic differences might result from a reduced activity of the *S. typhimurium* UmuC protein. Both the chromosomal *umuDC* genes are arranged in an operon. In *E. coli*, the genes overlap by one base pair, while in *S. typhimurium* they are separated by two nucleotides. One possible explanation for the results of Sedgwick *et al.*, is that the *S. typhimurium* UmuC protein is expressed at a higher level than its *E. coli* counterpart and by analogy to the *E. coli umuC* expression mutants, overexpression of the normal UmuC protein leads to the poorly mutable phenotype. To examine this possibility, we took advantage of the fact that both *umuC* genes have an identical *AgeI* restriction enzyme site located at the 5' end of the gene. (The only other conserved restriction enzyme site in the operon is *Clal* located within the *umuD* gene). This allowed us to construct chimeric genes that retained the translational linkage between the respective *umuD* and *umuC* genes but replaced 94% of the distal *umuC* sequences. It also permitted us to examine the ability of the Umu proteins to functionally complement *in cis* rather than in the *trans* situation described in previous studies. If the overexpression hypothesis were correct, we would have expected that the clone expressing *E. coli* UmuC from the *S. typhimurium* operon (pRW230) would exhibit low levels of mutagenesis similar to the native *S. typhimurium umuDC* operon (pRW224). However, replacing the *S. typhimurium* UmuC protein with that from *E. coli* resulted in a dramatic increase in the number of induced mutants). The converse was true with the reciprocal exchange event. The high level of mutagenesis seen with the native *E. coli* UmuDC proteins was dramatically reduced when the *S. typhimurium* UmuC protein replaced *E. coli* UmuC. These observations indicate that the *E. coli* and *S. typhimurium* UmuDC proteins can indeed substitute for each other and furthermore, that the structural *S. typhimurium* UmuC protein itself has a defect in its ability to promote mutagenesis functions.

We were interested in more precisely determining the location of the region(s) within the structural *S. typhimurium* UmuC protein responsible for the reduced mutability. Based upon the results with the chimeric plasmid pRW230, described above, it appears that it is not located in the first 25 amino acid residues but within the remaining 397 amino acids of *S. typhimurium* UmuC. Of these amino acids, 332 are identical in both *S. typhimurium* and *E. coli*, leaving just 65 residues as possible candidates for the phenotype. To identify the mutagenically important residues, we have taken advantage of the fact that the *E. coli* UmuDC proteins normally produce moderately high levels of UV-induced mutants. By introducing specific segments of the *S. typhimurium* UmuC protein into the reciprocal *E. coli* UmuC protein and assaying for a loss of mutagenesis-promoting activity we hoped to identify regions of *S. typhimurium* that were critical for mutagenic activity.

While one of the chimeric clones, pWK5, (which replaces residues 26-146), gave a non UV-mutable phenotype identical to the parental plasmid containing the entire *S. typhimurium* operon, another chimeric construct, pWK6, (expressing residues 147-395), also exhibited a

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

reduced ability to promote UV-mutagenesis, indicating that the defect might result from more than one change. This prompted us to subclone smaller fragments of the *S. typhimurium* *umuC* gene and assay the activity of each clone. The 26-146 stretch of *S. typhimurium* UmuC amino acid residues found in pWK5 was subcloned into two plasmids; one pWK7, (expressing residues 26-59) and the other pRW286 (expressing residues 62-146). Interestingly, plasmid pWK7 retained the same phenotype as the larger construct, pWK5, indicating that a mutation(s) within this region was responsible for the phenotype. This results in a complete loss of UV-induced mutagenesis. We would conclude that one (or more) of the 8 amino acid changes found in this short stretch of 34 residues severely curtails the activity of the chimeric UmuC protein's ability to promote mutagenesis. Of these eight altered residues, some appear more likely to cause the phenotype than others. For example, while the *S. typhimurium* UmuC differs from *E. coli* at residues Ser⁴¹, Ile⁴⁹, Glu⁵³ and Tyr⁵⁵ these residues are in fact found in other UmuC-like proteins, suggesting that they might be accommodated without any apparent loss of activity. The four remaining residues of Cys⁴⁰, Gln⁴⁶, Ala⁵⁰ and Pro⁵¹ are all unique amongst UmuC-like proteins studied to date and as such, appear to be good candidates for *S. typhimurium* UmuC's inability to promote UV-induced mutagenesis.

In contrast, the region between residues 62-146 (pRW286) were interchanged with only a modest loss of activity. Only 5 amino acids differ in this 85 amino acid region and it would appear that these changes are essentially neutral, at least for mutagenic activity.

Similar phenotypic effects were observed when we subcloned the 147-395 amino acid region found in pWK6. Plasmid, pRW254 containing residues 212-395 exhibited a phenotype similar to the larger chimeric construct indicating that a mutation within this region most likely caused the phenotype. In contrast, the chimeric clone expressing residues 147 to 209 (pRW304) was largely proficient at mutagenesis. Like the 5 changes in pRW286, it would seem that the 16 amino acid residues that differ between *S. typhimurium* and *E. coli* UmuC within this region are largely phenotypically silent.

In a previous study several loss of function mutants were located within the very COOH-terminal of the *E. coli* UmuC protein. The region between residues 396-422 (27 residues) contains 5 changes. When a chimeric protein was constructed that introduced these five changes into the *E. coli* UmuC protein, the number of UV-induced His⁺ mutants dropped significantly. This observation is therefore in good agreement with the suggestion that the integrity of the COOH-terminal of UmuC is important for mutagenic activity.

Based upon these findings, it appears that residues 62-209 from *S. typhimurium* UmuC can be interchanged with those from *E. coli* with only a modest loss of activity. Certain changes located within residues 212-395 and independently, 396-422, cause a significant reduction in activity, while those located within residues 26-59 caused the severest defects with regard to UV-induced mutagenesis and may be the major determinant of the phenotypically poor mutability of *S. typhimurium*.

Factors affecting the *in vivo* stability of the *E. coli* Umu proteins

We have previously determined that when fully derepressed, wild-type cells contain approximately 2400 UmuD molecules and 200 UmuC molecules per cell. Certain *recA* mutants constitutive for its coprotease activity (*recA**) have 3-5 fold higher steady-state levels of the Umu proteins. We have hypothesized that this increase might occur through a direct protein-protein interaction between the RecA* and Umu proteins that protects the Umu proteins from cellular degradation. To examine this possibility, we have constructed a novel *umuDC* mutation that substitutes the chromosomal *umuDC* genes with that of the *ermGT* gene encoding erythromycin resistance. The $\Delta(\textit{umuDC})596::\textit{ermGT}$ allele was then moved via generalized P1 transduction into *recA*⁺ and *recA730* backgrounds. Low copy plasmids encoding UmuD, UmuD' and UmuC were introduced into these two strains and the half-lives of the Umu protein analyzed. When expressed on its own UmuC is very unstable. When expressed simultaneously with UmuD, UmuC is partially stabilized. Much greater stabilization occurred however, in the presence of UmuD'. UmuC's half-life was longest in the *recA730* strain. Analysis revealed that UmuC is, however, not directly stabilized by RecA730, but rather was stabilized via RecA730's ability to stabilize UmuD' and thus indirectly stabilize UmuC. We are presently investigating other factors that might contribute to the stability of the Umu proteins *in vivo*.

Identification of a UmuC homolog in Archea bacteria

One of our long term goals is to identify mutagenesis proteins in higher organisms. To date, molecular analysis has revealed true *umu*-like genes in a variety of gram-negative enteric bacteria. Based upon DNA hybridization studies, a *umu* homolog has been reported in the gram-positive *Streptomyces coelicolor*, but the structure of these *umu*-like genes still awaits identification. Based upon their abundance at least in gram-negative enteric bacteria, and their apparent conservation throughout evolution, it is possible that the mutagenesis proteins in higher organisms are structurally, as well as functionally, related to the prokaryotic family of Umu-like proteins. To identify these proteins we have utilized degenerate oligonucleotides to regions that are 100% conserved of the previously identified UmuC-like proteins. Using this approach, we have identified a putative UmuC-like homolog in the Archea strain *Sulfolobus solfataricus*. Preliminary DNA sequence analysis revealed that this homolog is approximately 30% identical to UmuC over 200 amino acids residues. We are presently trying to characterize this new protein to determine its role in inducible mutagenesis.

Characterization of the MucAB proteins

To date, eight *umu*-like operons have been cloned and sequenced (including the three new *umu*-like genes that we have identified and are described above). Although these operons are structurally similar, the five previously identified *umu*-like operons all differ in their ability to enhance cellular and bacteriophage mutagenesis. Of these operons, the *mucAB* genes from the IncN group R-plasmid, pKM101, appears to be one of the most efficient at promoting mutagenesis. To determine the biochemical basis for the MucAB phenotype, we are attempting

to overproduce and purify the MucAB proteins and compare their biochemical activities to the purified *E. coli* Umu proteins. At the present time, we have overproduced and purified MucA'. We have also overproduced MucB proteins by coupling it to a strong T7 promoter. While most of the overproduced MucB was found in insoluble inclusion bodies we have used antibodies to MucB [that were kindly provided to us by H.Tanooka (Tokyo, Japan)], to detect significant quantities of MucB in soluble cell extracts. We are currently attempting to purify the soluble form of the MucB protein and are analyzing the ability of the MucB containing fractions to promote translesion DNA synthesis in a cell free system.

B. DNA REPAIR

In vitro functional studies of purified UV-DDB complex

This complex (p127+p41), first identified by us in primate cells, binds specifically to 6-4 pyrimidine dimers (photoproducts) arising in UV-irradiated DNA. Our collaborative studies with R. Wood have shown that our purified UV-DDB fraction is able to completely replace one of the fractions used in an *in vitro* nucleotide excision repair (NER) reconstituted system. Adding the purified protein complex to a mixture for *in vitro* repair synthesis led only to a two-fold stimulation of repair, suggesting that UV-DDB/XPE factor plays an accessory but not a core role in the NER process. Perhaps the factor has an auxiliary role in recognition (e.g., specific recognition of 6-4s with higher affinity than XPA-mediated recognition), or in the catalytic turnover of the incision/excision complex during repair of UV-irradiated DNA. Experiments are in progress to test whether the purified complex can restore full repair synthesis in extracts of XP-E fibroblasts that are defective in the binding activity, as shown *in vivo* by Linn's group, and to define the proper timing and positioning of UV-DDB in the process of NER.

Specific interaction of the UV-DDB p127 subunit with other components of NER

Using cell fractionation and specific antibodies, in our recent studies we have examined the nuclear distribution of UV-DDB/p127 and several other NER and DNA replication proteins before and after UV-treatment of TC7 cells. Following irradiation, UV-DDB/p127, as well as XPA, RP-A and PCNA, are translocated from low- to high-salt chromatin, suggesting that these proteins move to a tight association with damaged DNA. Contrary to XPA, only a small fraction of p127 is translocated to high-salt chromatin; this result follows the pattern of PCNA and RPA nuclear redistribution after UV. This implies that UV-DDB may have another cellular function(s) aside from its role in NER. Moreover, within this tight complex we find a specific interaction between UV-DDB/p127 and RPA *in vitro* and *in vivo*. RPA copurified with UV-DDB fractions which were isolated from mock or UV-treated primate cells. Within the same purified fractions we obtained coimmunoprecipitation of both proteins, and confirmed the significance of this interaction *in vivo* by coimmunoprecipitation from TC7 nuclear lysates. Under the same conditions, an *in vivo* specific protein interaction between UV-DDB and PCNA was not detected. A specific UV-DDB•RPA interaction results in enhanced binding to UV-damaged DNA. The purified RPA was added to an RPA-immunodepleted UV-DDB fraction

and binding to a probe was tested in the band-shift assay. The presence of both proteins enhanced binding activity about 5-fold and 3-fold compared to DDB or RPA binding alone, respectively. This enhanced binding is specific, can be competed out with a damaged competitor, and partially with an undamaged competitor, which correlates with our finding that RPA itself binds to dsDNA with a preference for UV-damaged over undamaged DNA. However, SSB does not affect DDB binding to damaged DNA (nor does BSA). The similarity of UV-DDB•RPA to XPA•RPA interactions, described by others, places UV-DDB in the damage-recognition step of NER.

SIGNIFICANCE

A. PROKARYOTIC MUTAGENESIS

The products of the *umuDC* genes are required for most inducible mutagenesis in *Escherichia coli*. Our studies indicate that related proteins are required for similar processes in other enterobacteria and even more evolutionarily evolved organisms. The biochemical role of these so-called mutagenesis proteins still remains to be elucidated. In an effort to address these problems the crystal structure of the *Escherichia coli* UmuD' mutagenesis protein was determined at 2.5Å resolution. The structure provides the first insights into the mechanism of the RecA mediated self cleavage reaction that this protein and a host of homologous proteins undergo upon induction of the SOS response. Analysis revealed that the structure is a very compact beta fold. Based upon the structure, we have hypothesized that in addition to forming a "molecular" dimer with itself, UmuD' can form an extended "polymer" structure that may play a role in positioning the UmuD' protein along the RecA-DNA filament.

Our studies with *Xenopus oocytes* and oocyte nuclear extracts revealed that while they are extremely efficient at both the repair of UV-photoproducts on double-stranded (ds) DNA and the replication of single-stranded (ss) to (ds) DNA, they are normally unable to replicate UV-irradiated ssM13 DNA containing cyclobutane pyrimidine dimers. The arrest was stage specific and abrogated in progesterone-matured oocytes. Moreover, the replication arrest was alleviated by injection into oocytes of mRNAs encoding the UmuD'C or MucA'B mutagenesis proteins. Our results may explain how cells stabilize repair/replication events on DNA with unrepairable lesions, and avoid DNA damage-induced apoptosis once irreversibly committed to S-phase. The results also suggest that a eukaryotic mutagenic pathway only becomes apparent after meiosis, and appear to reflect a remarkable conservation of prokaryotic to vertebrate replication machinery.

Thus, the molecular reagents developed from bacteria should provide information on functionally related proteins in higher organisms.

B. DNA REPAIR

In our studies on the mechanisms of DNA repair in mammalian cells, we have identified a DNA-binding activity, UV-DDB, which is specific for UV-induced 6-4 pyrimidine dimers. We

100
100

100

100

100

100

100

100

100

100

100

100

100

100

100

100

100

100

100

100

100

purified a UV-DDB protein complex from primate cells and isolated a cDNA encoding a 127 kDa component of the ~210 kDa complex. p127 is a structurally novel, evolutionarily conserved member of a family of proteins that may be involved in the recognition step of DNA repair. Because the UV-DDB activity is defective in at least some patients with the XP group E syndrome, we may be able to ascribe a specific repair function to this DNA-binding complex and identify a molecular defect in XP-E patients. This is of great interest because a biochemical defect in this sun-sensitive, cancer-prone disorder has not yet been defined. In addition, our finding that some mammalian tissues (brain, ovaries) do not express UV-DDB activity, and therefore may have a slower rate of removal of DNA lesions, sheds new light on the importance of DNA repair in rapidly proliferating tissues (e.g., skin) versus tissues that do not proliferate (e.g., brain). Furthermore, by examining UV-DDB activity in other DNA repair disorders, as well as in aging and cellular transformation, we should be able to obtain a much broader view of the role of DNA repair in the important life processes. In addition to its importance for human health, our discovery of the structural similarity of p127 to yet uncharacterized proteins from slime mold, *Drosophila* and rice is of consequence to our understanding of the evolutionary appearance of this DNA-binding protein — with respect to both its function and structure. Site-directed mutagenesis studies of p127, as well as identification of putative mutations in UV-DDB mRNAs isolated from XP-E patients who lack UV-DDB activity, should allow the definition of a novel, yet unrecognized, DNA-binding motif in the UV-DDB protein which interacts specifically with UV-damaged DNA.

PROPOSED COURSE

A. PROKARYOTIC MUTAGENESIS

While solving the crystal structure of the UmuD' protein has clearly provided great insights into the structure of the protein, much still remains to be determined. Perhaps the greatest interest is the role that the "polymer" structure plays in damage inducible mutagenesis. Using the calorimetric papillation assay that we have previously described and the yeast "two hybrid" system, we plan to identify mutants of *umuD'* that affect this interaction. Biochemical studies have previously shown that UmuD' interacts with both RecA and UmuC. While deletion of the first 20 N-terminal UmuD' residues severely reduced the ability of the protein to interact with RecA, there was still a weak interaction, suggesting that additional regions of the protein may also interact with RecA. Determining these sites, as well as those that interact with UmuC will provide us a better understanding of how these proteins interact to form the so-called "mutasome".

Our finding that the prokaryotic mutagenesis proteins can function in the heterologous *Xenopus laevis* system and thereby promote translesion DNA synthesis is especially exciting. We plan to extend these studies by using UV-irradiated derivatives of M13mp2 that can be used to distinguish if the bypass synthesis is mutagenic based upon a simple calorimetric assay. Analysis of the mutagenic spectra obtained in the presence and absence of the prokaryotic mutagenesis and comparison of the spectra found in the native prokaryotic environment should

provide additional evidence that the molecular mechanisms of mutagenesis are indeed remarkably conserved.

While we have been successful in partially purifying soluble MucB for biochemical analysis, it appears that the preparations contain a contaminating DNA polymerase activity. To circumvent these problems we are now attempting to purify MucB from an *Escherichia coli* strain deleted for DNA polymerase I and DNA polymerase II. Since the Muc proteins are known to act in a wide range of hosts, further characterization of the MucA'B proteins may directly provide us with an understanding of the mechanisms of mutagenesis in higher organisms.

We have previously identified several novel plasmid-encoded *umuC* mutants that have lost their ability to promote mutagenesis. At the present time, very little is known about UmuC function. Characterization of these mutants will therefore not only determine their biochemical defect, it should shed light on the activities of the wild-type UmuC protein. Of particular interest were the mutants that were located close to the COOH-terminal of UmuC. Using synthetic oligonucleotides we intend to generate a set of precise UmuC mutants that lack certain residues from the very end of UmuC. Based upon our present understanding of the mechanisms of mutagenesis, a ternary complex must form between UmuC, UmuD' and RecA proteins. It is possible that some (all) of these mutants are defective in these interactions. By using a combined approach of molecular, as well as genetic analyses, we might be able to gain insights into the potential structural domains of the UmuC protein as well as its functional activities.

While these studies outlined above will primarily focus on *Escherichia coli* and *Xenopus laevis*, we plan to continue characterization of the putative UmuC-like homolog that we have recently identified in *Sulfolobus solfataricus*. The DNA sequence of this clone will be extended to determine whether the homolog also expresses a UmuD-like homolog and whether the protein is, like UmuC, damage inducible.

Given that it does indeed appear that the molecular mechanisms of mutagenesis are conserved, we were interested to learn that another putative UmuC-like homolog has recently been identified as part of the *Saccharomyces cerevisiae* genome sequencing project. We plan to use standard yeast gene disruption techniques to determine if this homolog does indeed play a role in damage-induced mutagenesis.

B. DNA REPAIR

We will pursue the notion that the DDB complex may have a role in addition to recognition of UV lesions. One candidate role is transcription coupling, given our increasing awareness that stalled transcription induces DNA repair and that other XP proteins may serve as subunits of transcription factors. We shall also pursue the notion that the DDB complex may have a particularly high affinity for 6-4 photoproducts as opposed to other UV lesions, i.e., cyclobutane pyrimidine dimers. It may be that the XPA protein functions as a general damage-recognition protein, analogous to the *E.coli* Uvr A protein, but that other proteins serve as

accessory factors in the recognition of specific DNA conformations. Therefore, we plan to compare the affinity of DDB versus XPA with respect to binding to a probe containing only 6-4 lesions.

PUBLICATIONS

Aboussekhra A, Biggerstaff M, Shivji MK, Vilpo JA, Moncollin V, Podust VN, Protic M, Hubscher U, Egly JM, Wood RD. Mammalian DNA nucleotide excision repair reconstituted with purified protein components, *Cell* 1995;80:859-68.

Ennis DG, Levine AS, Koch WH, Woodgate R. Analysis of *recA* mutants with altered SOS functions, *Mutat Res* 1995;336:39-48.

Kuleava OI, Wootton JC, Levine AS, Woodgate R. Characterization of the *umu*-complementing operon from R391, *J Bacteriol* 1995;177:2737-2743.

Zernik-Koback M, Szumiel I, Levine AS, Dixon K. Analysis of mutagenesis in UV-sensitive mouse lymphoma L5178Y-R cells with a polyomavirus-based shuttle vector. *Mutat Res* 1995;344:31-39.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER

Z01 HD 2400-04 PB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

The role of subclinical infection and cytokines in preterm parturition

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

Roberto Romero, M.D. Head PRB, NICHD

Others:

Ricardo Gomez, M.D. Visiting Fellow PRB, NICHD

Fabio Ghezzi, M.D. Visiting Fellow PRB, NICHD

Jose Cohen, M.D. Visiting Fellow PRB, NICHD

Jorge Tolosa, M.D. Senior Staff Fellow PRB, NICHD

Moshe Mazor, M.D. Visiting Fellow PRB, NICHD

COOPERATING UNITS (if any)

Department of Obstetrics and Gynecology, Wayne State University (Stanley Berry, MD, and Ruben Quintero, MD)

LAB/BRANCH

Perinatology Research Branch

SECTION

INSTITUTE AND LOCATION

NICHD, Washington, DC 20007

TOTAL STAFF YEARS:

4

PROFESSIONAL:

OTHER:

CHECK APPROPRIATE BOX(ES)

☒ (a) Human subjects ☒ (b) Human tissues ☐ © Neither

☒ (a1) Minors

☒ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The laboratory concluded the following studies:

1. **Interleukin-6 concentrations in umbilical cord plasma identify infants at risk for the development of white matter lesions associated with periventricular leukomalacia (PVL):** PVL is a common white matter lesion of the central nervous system of preterm neonates and a major risk factor for cerebral palsy. A series of studies have been designed to test the hypothesis of whether cytokine-mediated neural injury is responsible for this condition. Investigators of the Branch found that IL-6 determinations in umbilical cord can identify infants at risk for PVL-associated lesions.

2. **Amniotic fluid interleukin-6 as a marker for subclinical infection limited to the chorio-amniotic space:** Previous studies conducted by the Branch have demonstrated that amniotic fluid concentrations of interleukin-6 can serve as a sensitive and specific method for the diagnosis of intraamniotic infection. We have now established that patients with subclinical extra-amniotic infection localized to the space between amnion and chorion have also elevated interleukin-6 concentrations. This observation has diagnostic and therapeutic implications. Patients with elevated amniotic fluid IL-6 and negative amniotic fluid cultures have been considered "false negative cases" of intra-amniotic infection. However, they most likely represent patients with early stages of intrauterine infection which are undetected by amniotic fluid culture. Therapeutically, this sub-group of patients represents the ideal population in which to test whether antibiotic treatment can reduce perinatal morbidity and mortality and delay preterm delivery. Previous attempts to treat intra-amniotic infection with antibiotics have been unsuccessful because the infection is too far advanced.

3. **In utero activation of the fetal immune system during preterm parturition:** During the birth process the fetus emerges from the normally sterile intrauterine environment to the non-sterile extrauterine environment. To determine if parturition is associated with changes in fetal immune adaptation, we performed fetal blood sampling in fetuses in premature labor. Fetuses who delivered preterm (< 72 hours of the procedure) had a higher median number of positive cells for CD11c, CD13, CD15 and CD67 (markers for neutrophil and/or monocyte activation) but not CD14 (receptor for the LPS-LPS binding protein) and CD63 (platelet activation) than those delivering at term. Our findings suggest that preterm parturition is associated with activation of the fetal monocyte-neutrophil systems which are the first cellular line of defense against infection. Therefore, important changes in immune adaptation occur in the human fetus during labor.

PROJECT TITLE: The role of subclinical infection and cytokines in preterm parturition

PROJECT DESCRIPTION

The long term goal of this project is the prevention of preterm birth, the leading cause of perinatal morbidity and mortality worldwide. The short term objective is to increase understanding of the basic mechanisms of disease responsible for preterm parturition.

The specific objectives for this year were:

1. To determine the clinical significance of microbial invasion of the amniotic cavity with genital Mycoplasmas.
2. To examine the relationship between amniotic fluid interleukin-6 concentrations and infection in the extra-amniotic space.
3. To determine if the fetal monocyte-neutrophil system is activated during spontaneous preterm parturition.
4. To establish if changes in impedance to blood flow occur in the fetal cerebral circulation during preterm parturition.
5. To examine the relationship between umbilical cord cytokine concentrations and the subsequent development of periventricular leukomalacia.

Methods employed:

1. Transabdominal amniocentesis to retrieve amniotic fluid.
2. Cordocentesis to obtain human fetal blood.
3. Sensitive and specific immunoassays for macrophage derived cytokines.
4. Sorting and counting of fetal cell using FACS with specific monoclonal antibodies.
5. Doppler examination of the fetal circulation with ultrasound.

Progress

1. The clinical significance of microbial invasion with ureaplasma urealyticum in the midtrimester of pregnancy: Amniotic fluid is

normally sterile and the isolation of any microorganisms from amniotic fluid retrieved by transabdominal amniocentesis is an abnormal finding referred to as "microbial invasion of the amniotic cavity". We conducted a study to determine the prevalence and frequency of microbial invasions with genital Mycoplasmas. Subclinical microbial invasion was found in 2.8% of cases and the frequency of adverse pregnancy outcome was higher in patients with ureaplasma urealyticum isolated from amniotic fluid than in those with sterile fluid. The adverse pregnancy outcomes included preterm premature rupture of membranes, preterm delivery and preterm labor. Our observations justify further studies to determine if routine culture for Mycoplasmas should be performed in all women undergoing genetic amniocentesis and if treatment with antibiotics may be effective in improving pregnancy outcome.

2. Amniotic fluid interleukin-6 as a marker for subclinical infection limited to the chorio-amniotic space: Previous studies conducted by the Branch have demonstrated that amniotic fluid concentrations of interleukin-6 can serve as a sensitive and specific method for the diagnosis of intraamniotic infection. However, we had also noted that patients with preterm labor and elevated interleukin-6 concentrations are refractory to tocolysis and have histologic chorioamnionitis even in the absence of demonstrable intraamniotic infection. It has been hypothesized that the elevation of amniotic fluid IL-6 in these patients is due to an extra-amniotic space. Chorio-amniotic inflammation would lead to an increase in amniotic fluid IL-6 concentrations even though amniotic fluid cultures are negative. To test this hypothesis, amniotic fluid IL-6 concentrations were determined in a group of women in whom microbial cultures were obtained from both amniotic fluid and the chorio-amniotic space. We found that patients with microbial invasion limited to the chorio-amniotic space had higher amniotic fluid concentrations of IL-6 than those with negative microbial cultures of the membranes. These findings support the view that a sub-group of patients with preterm labor and elevated amniotic fluid IL-6 contractions have an intrauterine infection limited to the extra-amniotic space. This observation has diagnostic and therapeutic implications. Patients with elevated amniotic fluid IL-6 and negative amniotic fluid cultures have been considered "false negative cases" of intra-amniotic infection. However, they most likely represent patients with early stages of intrauterine infection which are undetected by amniotic fluid culture. From a therapeutic point of view, this sub-group of patients may represent the most ideal population for antibiotic treatment. Previous efforts of antibiotic treatment in patients who had proven intraamniotic infection has been unsuccessful.

3. In utero activation of the fetal immune system during preterm parturition: During the birth process the fetus emerges from the normally sterile intrauterine environment to the non-sterile extrauterine environment. Little is known of the changes that the human fetal immune system undergoes during this transitional period. In collaboration with Dr. Stanley Berry of Wayne State University, fetal blood sampling was performed in patients with

preterm labor to determine whether activation of the human fetal neutrophil and monocyte systems occur during preterm labor and delivery. Fetuses who delivered preterm (< 72 hours of the procedure) had a higher median number of positive cells for CD11c, CD13, CD15 and CD67 (markers for neutrophil and/or monocyte activation) but not CD14 (receptor for the LPS-LPS binding protein) and CD 63 (platelet activation) than those delivering at term. These results remained significant after correcting for the administration of antenatal steroids or for the presence of intra-amniotic infection. Our findings suggest that preterm parturition is associated with activation of the fetal monocyte-neutrophil systems which are the first cellular line of defense against infection. Therefore, important changes in immune adaptation occur in the human fetus during labor.

4. Doppler velocimetry of the fetal middle cerebral artery in patients with preterm labor and intact membranes: Intrapartum Doppler velocimetry studies of fetal cerebral circulation in human fetuses have shown that labor at term is associated with a 40% decrease in the impedance of flow in the middle cerebral artery (MCA). This observation has been interpreted as a response to a decreasing fetal pO_2 during labor. Previous studies have reported that preterm labor is associated with changes in impedance of flow in the umbilical artery. However, no information is available as to whether or not changes in cerebral vascular impedance of flow occur in this condition. To examine this question, 194 fetuses in preterm labor with intact membranes underwent Doppler examination of the middle cerebral artery. Patients who delivered within 24 hours of examination had a lower pulsatility index corrected for gestational age than those who delivered at term. Our observations suggest that a decreased impedance of blood flow in the cerebral circulation occurs in both term and preterm parturition. The most likely explanation for these findings is that uterine contractions during labor result in a relative decrease in fetal pO_2 which in turn leads to cerebral vasodilation. This work provides new insights into the physiologic adaptation to imminent delivery.

Intrauterine Infection, Cytokines and Cerebral Palsy: Periventricular leukomalacia (PVL), a cerebral lesion characterized by foci of necrosis in the white matter near the lateral ventricles, is found in 7 to 26% of babies with birth-weight under 1,500g. Forty-seven to 80% of these infants developed abnormalities of neuromuscular tone and posture and often cognitive impairment. Although the etiology of this lesion has not been defined, clinical risk factors include prematurity, asphyxia, respiratory distress requiring mechanical ventilation and septicemia. Recently, the role of infection in the pathogenesis of PVL has become a major focus of attention because neonates with sepsis or those born to mothers with documented infection are at increased risk for PVL. A role for cytokines and specifically tumor necrosis factor (TNF α) has been recently postulated in the pathophysiology of PVL. To explore this hypothesis, the Branch conducted a study to examine the relationship between umbilical cord plasma TNF α , interleukin-6 (IL-6) and IL-1 receptor antagonist (IL-1ra) and the occurrence of PVL

within the first 3 days of life in 172 preterm births. Plasma concentrations of IL-6, but not of TNF, IL-1 beta and IL-1ra, were significantly higher in neonates with PVL-associated lesions than in those without these lesions. Moreover, we found, after correcting for other confounding variables (odds ratio: 6.2), that the independent predictor of neonatal PVL-associated lesions. Our findings support the hypothesis that intrauterine fetal infection may lead to activation of the cytokine network, white matter damage and PVL. The lack of a relationship between umbilical cord plasma TNF and subsequent development of PVL does not exclude a role for this cytokine in the development of PVL. This finding may be due to the differential kinetic pattern of plasma cytokine concentrations in response to infection. In animals as well as in human volunteers the administration of lipopolysaccharide (LPS) leads to an increase in plasma concentrations of TNF, followed by IL-1B and then IL-6. However, the increase in plasma IL-6 concentrations occurs at a time when TNF concentrations have returned to nearly normal or non-detectable levels. Therefore, it is possible that an intrauterine infection stimulated fetal TNF production, which in turn led to brain white matter damage, although evidence of TNF excess may not be detectable at the time of birth. An elevated IL-6 level in the umbilical cord plasma may be a marker of previous stimulation of the feto-placental immune system by microbial products. An important clinical implication of our study is that IL-6 determinations in umbilical cord plasma represent a simple means of identifying neonates at risk for PVL and subsequent cerebral palsy.

Significance

1) Significance of the problem: Prematurity is the leading cause of perinatal morbidity and mortality worldwide. At least 75% of perinatal deaths not due to congenital anomalies are associated with prematurity. Moreover, survivors are at an increased risk for long term neurologic and developmental handicaps. Infants born at < 900g cost society more than they are likely to earn in their life time. Thus, prematurity is considered the single most important problem in perinatal medicine today. Despite the availability of tocolytic therapy (pharmacologic agents to stop uterine contractions) and government funded programs to improve maternal nutrition and prenatal care, the preterm delivery rate in this country has remained essentially unchanged at approximately 9%. This represents at least 180,000 preterm births per year.

Studies performed by our group indicate that at least 25% of all preterm births are associated with subclinical microbial invasion of the amniotic cavity. Moreover, women with subclinical intrauterine infections are refractory to tocolysis and their neonates are at higher risk for perinatal complications. Thus, this group of patients contributes significantly to perinatal morbidity and mortality associated with preterm birth. The studies conducted by this laboratory are designed to define the role of microbial invasion in preterm birth, to elucidate the cellular and molecular mechanisms responsible for preterm birth, and to develop diagnostic

and therapeutic modalities for the treatment of preterm labor and prevention of preterm birth.

2) Significance of specific accomplishments:

1. The clinical significance of microbial invasion with ureaplasma urealyticum in the midtrimester of pregnancy: The finding that subclinical microbial invasion of the amniotic cavity occurs in normal women and that it is a risk factor for poor pregnancy outcome has major clinical implications. Genetic amniocentesis is the most frequently performed invasive test for prenatal diagnosis. Our data suggest that routine amniotic fluid culture for genital mycoplasmas should be considered in all women undergoing this procedure.

2. Elevated amniotic fluid interleukin-6 identifies patients with subclinical intrauterine infection limited to the chorio-amniotic space: Based upon previous studies, the Perinatology Branch has established that patients with an elevated amniotic fluid IL-6 concentration but a negative amniotic fluid culture for microorganisms are at risk for preterm delivery. The results generated this year indicate that a fraction of these patients have a subclinical infection limited to the extra-amniotic space (chorio-amniotic space). Since the microorganisms have not invaded the amniotic cavity, amniotic fluid cultures are negative. These observations have biological, diagnostic and therapeutic implications. First, subclinical intrauterine infection may be present even in the absence of positive amniotic fluid cultures. Second, these patients can be identified by measuring a cytokine in the amniotic fluid: IL-6. Third, although antibiotic administration has not been beneficial in the treatment of patients with proven intra-amniotic infection (positive amniotic fluid culture), they may benefit the sub-group with early extra-amniotic infections. The next step is to conduct a randomized clinical trial to address this question.

3. In utero activation of the fetal immune system during preterm parturition: Our results indicate that the effector limb of the fetal immune system (monocyte-neutrophil) is activated during the process of spontaneous preterm parturition. This activation is present even in the absence of infection and probably represents immune adaptation as the fetus emerges from a sterile intrauterine environment to the non-sterile extrauterine environment.

4. Doppler velocimetry of the fetal middle cerebral artery in patients with preterm labor and intact membranes and impending preterm delivery: The findings that fetuses in preterm labor who delivered within 24 hours of examination had a lower pulsatility index in the middle cerebral artery than those who delivered at term suggest that parturition leads to a decreased impedance of blood flow in the fetal cerebral circulation. This phenomenon is present in both preterm and term labor and we believe that it represents a hemodynamic adaptation of the fetus to protect its brain against the decrease PO_2 associated with intermittent uterine

contractions.

5. Intrauterine Infection, Cytokines and Cerebral Palsy: Prematurity is the leading identifiable cause of cerebral palsy. Periventricular leukomalacia (PAL), a cerebral lesion characterized by foci of necrosis in the white matter near the lateral ventricles, is found in 7 to 26% of babies with birth-weight under 1000g and is a major risk factor for cerebral palsy. The Branch has demonstrated that there is a relationship between subclinical infection and the subsequent development of PAL and that neonates at risk for PAL can be identified by an elevated concentration of IL-6 in umbilical cord blood.

Mazor M, Chaim W, Horowitz S, Romero R, Glezerman M. The biomolecular mechanisms of preterm labor in women with intrauterine infection. *Isr J Med Sci* 1994; 30:317-22

Mazor M, Chaim W, Romero R. Preterm labor syndrome. *Harefuah* 1995; 128:111-16

Romero R, Munoz H, Gomez R, Sherer DM, Ghezzi F, Ghidini A, Alfi O, DeVore G, Randolph L. Two-thirds of spontaneous abortion/fetal deaths after genetic midtrimester amniocentesis are the result of a pre-existing subclinical inflammatory process of the amniotic cavity. *Am J Obstet Gynec* 1995; 172:261

Yoon B, Romero R, Yang SH, Jun JK, Choi JH, Kim IO, Syn HC, Kim SW. Interleukin-6 concentrations in umbilical cord plasma identify infants at risk for the development of white matter brain lesions associated with periventricular leukomalacia. *Am J Obstet Gynec* (in press)

Andrews WW, Hauth JC, Goldenberg RL, Gomez R, Romero R, Cassell GH. Amniotic fluid interleukin-6: Correlation with upper genital tract microbial colonization and gestational age in women delivered following spontaneous labor versus indicated delivery. *Am J Obstet Gynecol* 1995; 173(2):606-12.

Cohen J, Ghezzi F, Romero R, Ghidini A, Mazor M, Tolosa J, Goncalvez L, Gomez R. GRO alpha in the fetal maternal and amniotic fluid compartments during pregnancy and parturition. *Am J Repr Immun* (in press).

Dombrowski MP, Bottoms SF, Saleh AA, Hurd WW, Romero R. Third stage of labor: Analysis of duration and clinical practice. *Am J Obstet Gynecol* 1995; 172:1279-84.

Ghezzi F, Ghidini A, Romero R, Gomez R, Galasso M, Sherer DM, Cohen J, Treadwell MC, Cotton DB. Doppler velocimetry of the fetal middle cerebral artery in patients with preterm labor and intact membranes; A prospective study of 194 consecutive cases. *J Ultrasound Med* 1995; 14:361-66.

Gomez R, Ghezzi F, Romero R, Munoz H, Tolosa J, Rojas I. Premature labor and intraamniotic infection: Clinical aspects and role of cytokines in diagnosis and pathophysiology. *Clin Perinat* 1995; 22:281-342.

Gomez R, Ghezzi F, Romero R, Munoz H, Sherer DM, Rojas I. Premature labor and intraamniotic infection: Clinical aspects and role of the cytokines in diagnosis and pathophysiology. *Clinics in Perinatology* 1995; 22:281-342.

Gomez R, Romero R, Galasso M, Behnke E, Insunza A, Cotton DB. The value of amniotic fluid interleukin-6, white blood cell count and Gram stain in the detection of microbial invasion of the amniotic cavity in patients at term. *Am J Reprod Immunol* 1994; 32:200-10.

Horowitz S, Mazor M, Romero R, Horiwitz J, Glezerman M. Infection of the amniotic cavity with *Ureaplasma urealyticum* in the midtrimester of pregnancy. *J Reprod Med* 1995; 40:375-79.

Jarvis JN, Deng L, Berry SM, Romero R, Moore H. Fetal cytokine expression in utero detected by reverse transcriptase polymerase chain reaction. *Ped Research* 1995; 37:450-4.

Mazor M, Chaim W, Romero R. Preterm labor--a pathological process. *Isr J Obstet Gynecol* 1994; 5:92-96.

Mazor M, Cohen J, Romero R, Ghezzi F, Tolosa J, Gomez R. Cytokines and preterm labor. *Fetal Maternal Med Review* (in press).

Petraglia FP, Aguzzoli L, Florio P, Baumann P, Genazzani AD, DeCarlo, C, Romero R. Maternal plasma and placental immunoreactive corticotrophin-releasing factor concentrations in infection-associated term and preterm delivery. *Placenta* 1995; 16(2):157-64.

Petraglia F, DeVita D, Gallinelli A, Aguzzoli L, Gennazzani AR, Romero R, Woodruff TK. Abnormal concentration of maternal serum activin-A in gestational diseases. *J Clin Endocrin Metab* 1995; 80:558-61.

Romero R, Munoz H, Gomez R, Galasso M, Sherer D, Cotton D, Mitchell M. Does infection cause premature labor and delivery? *Semin Reprod Endocrinol* 1994; 12:227-48

Waring PM, Romero R, Laham N, Gomez R, Rice GE. Leukemia inhibitory factor: Association with intra-amniotic infection. *Am J Obstet Gynecol* 1994; 171:1335-41.

Yoon BH, Romero R, Kim CJ, Jun JK, Kim KS, Gomez R, Koo JN, Lee JH, Choi JH, Syn HC, Kim SW. Amniotic fluid interleukin-6: A useful test for antenatal diagnosis of acute inflammatory lesions of preterm placenta and prediction of perinatal morbidity. *Am J Obstet Gynecol* 1995; 172:960-70.

DEPARTMENT OF HEALTH AND HUMAN SERVICES - PUBLIC HEALTH SERVICE
NOTICE OF INTRAMURAL RESEARCH PROJECT

PROJECT NUMBER
 Z01 HD 2401-03 PB

PERIOD COVERED

October 1, 1994 to September 30, 1995

TITLE OF PROJECT (80 characters or less. Title must fit on one line between the borders.)

Prenatal Diagnosis of Congenital Anomalies

PRINCIPAL INVESTIGATOR (List other professional personnel below the Principal Investigator.) (Name, title, laboratory, and institute affiliation)

Roberto Romero, M.D. Head PRB, NICHD

Others:

Jorge Tolosa, M.D.	Senior Staff Fellow	PRB, NICHD
Ricardo Gomez, M.D.	Visiting Fellow	PRB, NICHD
Fabio Ghezzi, M.D.	Visiting Fellow	PRB, NICHD
Jose Cohen, M.D.	Visiting Fellow	PRB, NICHD
Moshe Mazor, M.D.	Visiting Fellow	PRB, NICHD

COOPERATING UNITS (if any)

Department of Obstetrics and Gynecology, Wayne State University (Stanley Berry, MD, and Ruben Quintero, MD)

LAB/BRANCH

Perinatology Research Branch

SECTION

INSTITUTE AND LOCATION

NICHD, Washington, DC 20007

TOTAL STAFF YEARS:

4

PROFESSIONAL:

OTHER:

CHECK APPROPRIATE BOX(ES)

☒ (a) Human subjects ☒ (b) Human tissues ☐ © Neither
☒ (a1) Minors
☒ (a2) Interviews

SUMMARY OF WORK (Use standard unreduced type. Do not exceed the space provided.)

The focus of this project is to improve the diagnosis and treatment of fetal disease. Focus is placed on the prenatal diagnosis of congenital anomalies with non-invasive methods (e.g., high resolution ultrasound and color Doppler flow mapping) and invasive methods.

The Branch pioneered the utilization of thin-gauge endoscopy in the differential diagnosis of fetal obstructive uropathy. Fetuses with sonographic findings of lower obstructive uropathy are a diagnostic and therapeutic challenge. Ultrasound imaging alone cannot establish the cause of the obstruction and treatment with percutaneous vesico-amniotic shunts has a complication rate of 25%. We inserted a fiber optic endoscope through the lumen of the needle (or trocar) placed into the fetal bladder and visualized the bladder neck and ureteral orifices. Eleven fetal cystoscopies were performed successfully. Ureteral webs were noted in two fetuses. A catheter was passed from the fetal bladder into the amniotic cavity in two cases to treat the obstruction. Ureteral probing with a flexible catheter was sufficient to treat the obstruction in other cases. This impressive technological achievement is a landmark in fetal medicine and surgery.

High resolution ultrasound was used to establish the prenatal diagnosis of congenital anomalies in the cardiovascular system. Color Doppler flow mapping has improved the accuracy of prenatal diagnosis of several other conditions which could only be suspected by gray-scale sonography. During the past year, the investigators of the Branch reported the first diagnosis of agenesis of the right and left portal veins. The Branch also described a diagnostic approach for the identification of coarctation of the aorta with the aid of color-flow Doppler ultrasound and identified an aneurism of the ductus arteriosus.

PROJECT TITLE: Prenatal Diagnosis of Congenital Anomalies

PROJECT DESCRIPTION:

The Perinatology Research Branch continues its pioneering work in the development of fetal endoscopic surgery in humans. Over the past two years, the Branch, in collaboration with Dr. Quintero from Wayne State University, developed techniques to conduct in utero fetal cystoscopy and endoscopic surgery of the fetal urinary tract.

SPECIFIC AIMS:

1. To perform endoscopic intrauterine fetal surgery in some congenital anomalies amenable to correction.
2. To further the knowledge of prenatal diagnosis of congenital anomalies.

METHODS EMPLOYED:

1. High resolution ultrasound and color Doppler flow mapping are two non-invasive techniques already utilized to establish prenatal diagnosis of conditions difficult or impossible to recognize in the past.

2. Transabdominal thin-gauge embryofetoscopy is a technique used to gain access to the fetus and the intrauterine environment. This procedure allows visualization of the surface anatomy of the human embryo and fetus, as well as insertion of surgical instruments into the amniotic cavity under direct visualization. Fetal surgery can thus be performed for the correction of some congenital anomalies.

PROGRESS REPORT:

1. Endoscopic intrauterine surgery in the human fetus:

The Perinatology Research Branch continues its pioneering work in the development of fetal endoscopic surgery in humans. Over the past two years, the Branch, in collaboration with Dr. Quintero from Wayne State University, developed techniques to conduct in utero fetal cystoscopy and endoscopic surgery of the fetal urinary tract.

Percutaneous Fetal Cystoscopy: Congenital obstruction of the fetal urinary tract is a common and serious congenital anomaly which can lead to perinatal death or permanent chronic renal failure. Lower urinary tract obstruction can be caused by posterior urethral valves, urethral atresia, vesico-urethral reflux and other causes of bladder outlet obstruction. Obstruction of urine outflow can lead to dilatation of the bladder (megacystis), retrograde dilatation of the ureters (megaureters) and hydronephrosis. If obstruction takes place early in fetal life, permanent renal damage may occur (renal dysplasia). Renal insufficiency results in

decreased or absent urine production as reflected by decreased amniotic fluid volume (oligohydramnios). Since fetal lung development is critically dependent upon amniotic fluid volume, oligohydramnios in early pregnancy can lead to pulmonary hypoplasia, the most common cause of death in newborns with congenital obstruction of the urinary tract. The prenatal diagnosis of lower urinary tract obstruction can be made with ultrasound but a precise diagnosis of the cause may not be possible. On the other hand, fetal cystoscopy allows direct examination of the ureteral orifices, urethra and bladder and assists in making a precise etiologic diagnosis.

Last year we reported the feasibility of performing fetal cystoscopy for diagnosis and its potential value in treatment. This year, we have used the procedure to assess eleven fetuses suspected to have lower obstructive uropathy by ultrasound examination. Fetal cystoscopy allowed the identification of bladder mucosal edema, hemorrhage and trabeculation which cannot be diagnosed with ultrasound. Moreover, endoscopic examination made possible the identification of ureteral webs at the uterovesical junction in two fetuses and dilated ureteral orifices suggesting vesicoureteral reflux in two fetuses. In one case, in which the status of the urethra could not be ascertained by ultrasound, fetal cystoscopy identified the structure to be dilated due to obstruction and therefore excluded urethral agenesis as the cause. In another case, the combination of ultrasound and endoscopic finding allowed the prenatal diagnosis of "megacystis-microcolon syndrome", a congenital non-obstructive cause of megacystis.

Of significance is that in two cases we performed, for the first time, a transurethral vesico-amniotic shunting. Under endoscopic guidance, a 5-French double-pig-tailed catheter was passed over a wire guide into the fetal bladder and through the urethra allowing fetal urine to flow from the bladder into the amniotic cavity and thus bypassing the obstruction. This approach is an alternative to the standard treatment of percutaneous placement of a vesico-amniotic shunt which has significant limitations. First, it is a palliative measure that defers the final treatment of the obstructions until the birth of the child. Endoscopic directed ablation of posterior urethral valves could result in permanent correction. Second, vesico-amniotic shunts become obstructed or displaced in up to 25% of cases, requiring additional intervention for replacement of the shunt. Such problems are less likely to occur with endoscopically directed procedures.

The advantages of percutaneous fetal cystoscopy can be realized without posing additional risk to the standard assessment of fetuses with lower obstructive uropathy which requires vesicocentesis (placement of a needle into the fetal bladder to drain urine and obtain specimens for analysis). Although a slightly wider needle is required for fetal cystoscopy (18 versus 22 gauge) newer fiber optic endoscopes may soon be able to be introduced through smaller caliber needles. Percutaneous fetal cystoscopy at the time of shunt placement uses a similar trocar diameter to that

of the shunt. The use of endoscopy will improve the understanding of causes of lower urinary tract obstruction in the human fetus, enhance prognostic assessment and allow the introduction of novel therapeutic modalities to improve the outcome for the sick fetus.

2. Prenatal diagnosis of congenital anomalies:

Aneurysm of the ductus arteriosus: The ductus arteriosus is a fetal vessel which permits physiologic shunting of oxygenated blood from the pulmonary artery to the aorta during intrauterine life. Constriction of the ducts during intrauterine life (i.e. induced after administration of indomethacin for the treatment of preterm labor) may lead to right ventricular dysfunction while complete closure of the ducts may result in intrauterine death. Ultrasound allows detailed examination of the ductus arteriosus during intrauterine life. We recently reported the successful diagnosis of a ductus arteriosus aneurysm in a fetus who had two previous normal ultrasound examinations at 20 and 29 weeks of gestation and a normal karyotype. At 39 weeks of gestation, an enlarged dilated vascular structure measuring 15mm in diameter was noted to arise from the bifurcation of the main pulmonary artery and to end in the descending aorta. The branches of the pulmonary artery were of normal dimensions. Color flow mapping of the structure demonstrated turbulent flow from the pulmonary artery to the descending aorta. The diagnosis of aneurysm of the ductus arteriosus was confirmed with echocardiography at birth. Doppler studies demonstrated bi-directional flow of the shunt predominantly right to left through the ducts. At 4 weeks of age magnetic resonance imaging documented closure of the ductus arteriosus. The diagnosis of these conditions is important because significant complications have been reported in affected neonates (rupture, erosion into a bronchus or esophagus and thromboembolism). Although surgical treatment has been advocated, this case demonstrates that spontaneous closure may occur. In addition, our case demonstrates that this anomaly may be absent in early intrauterine life and develop in the third trimester.

Fetal cardiac fibroma: Fetal cardiac tumors are serious conditions associated with a poor prognosis. A recent review documents intrauterine death in 57% of cases. The most common tumors are rhabdomyomas, teratomas and fibromas. We recently diagnosed a cardiac tumor in a fetus in the third trimester, after two previous ultrasound examination at 13 and 21 weeks had been reported as normal. Fetal echocardiography demonstrated a 2.7 x 2.5 cm echogenic mass originating from the free wall of the right ventricle and mild tricuspid regurgitation. After birth, enucleation of the tumor was performed and histologic examination demonstrated that this was consistent with a cardiac fibroma. Our case demonstrates the potential good prognosis of this condition which may be attributed to late onset of the tumor, its sessile base and its location away from the outflow tracts of the ventricles. In the only other reported case in the literature, the parents elected termination of pregnancy.

The use of color Doppler in the prenatal diagnosis of coarctation of the aorta: Coarctation of the aorta represents 6-9% of all congenital cardiovascular abnormalities. Prenatal diagnosis of this condition is rare because visualization of aortic narrowing with ultrasound in a human fetus is extremely difficult. Indeed, the condition is generally suspected by the presence of indirect signs (i.e., enlarged right ventricle and ventricular size disproportion). We recently diagnosed coarctation of the aorta with the use of color Doppler imaging (CDI) in a 38 week fetus who had a perimembranous ventricular septal defect. Views of the outflow tracts demonstrated a normal main trunk of the pulmonary artery but an abnormal aortic isthmus. The markedly hypoplastic transverse aortic arch was demonstrated by CDI proximal to the insertion of the ductus arteriosus. The narrow isthmus suggested a tubular coarctation. Prenatal diagnosis was confirmed post-natally. Prostaglandin E was administered to maintain the ductus open until corrective surgery was performed (consisting of amplification of the aorta with a Teflon patch, resection of the area of coarctation with end to end anastomosis and banding of the pulmonary artery). The infant is asymptomatic 8 months after birth.

The Perinatology Branch continues its pioneering work in the development of fetal endoscopic surgery in humans. Last year the Branch reported the first successful endoscopic surgical procedure performed on a human fetus: ligation of the umbilical cord of an acardiac twin. This year we have developed techniques to conduct in utero fetal cystoscopy and surgery of the fetal urinary tract.

SIGNIFICANCE

Congenital anomalies are the second leading cause of perinatal mortality and are a major focus of research activity for our team. The Perinatology Branch has developed a novel method to access the utero-fetal environment: embryofetoscopy. This technique has allowed diagnosis of conditions beyond the resolution of ultrasound (e.g., polydactyly in a first trimester fetus). More importantly, we have proved its feasibility in the in-utero treatment of some congenital anomalies amenable to surgical repair. We foresee that future applications will also include access to the embryonic and fetal circulation for therapeutic purposes (e.g., gene therapy). This procedure, though more invasive than amniocentesis and percutaneous fetal blood sampling, has the advantage of avoiding the risks of hysterotomy to correct diseases of the human fetus amenable to surgical treatment.

Ultrasound studies of the fetus and Doppler interrogation of the fetal hemodynamic circulation, including color Doppler flow mapping, have been of value in expanding the number of congenital anomalies which can be prenatally diagnosed, as well as improving the accuracy of gray scale ultrasound.

PROPOSED COURSE

We plan to conduct the following studies:

1. Expand the diagnostic and therapeutic indications for embryofetoscopy in the human pregnancy.
2. Determine the value of ultrasound examination, fetal echocardiography, and Doppler flow studies of the maternal and fetal peripheral vascular system in the prediction of preterm delivery in patients at risk for this complication.
3. Determine the value of sonographically-derived fetal biometry in the second trimester to predict the growth potential of the fetus and evaluate the clinical significance of deviation from the model.

Goncalves LF, Sherer DM, Silva M, Ghezzi F, Treadwell MC, Romero R. Prenatal sonographic demonstration of agenesis of the right and left portal veins and associated intrahepatic postsystemic shunts. *J Ultrasound Med* (in press).

Sherer DM, Romero R, Silva M, DeVore G, Goncalves LF, Ghezzi F, Cohen J, DeVore GR. Prenatal diagnosis of coarctation of the aorta with Color Doppler Imaging. *Acta Obstet Gynec* (in press)

Treadwell MC, Sherer DM, Sacks AJ, Ghezzi F, Romero R. Resolution of nonimmune hydrops secondary to fetal hypothyroidism following cordocentesis and adjusted medical therapy. *Obstet Gynecol* (in press).

Lauria MR, Gonik B, Romero R. Pulmonary hypoplasia: Pathogenesis, diagnosis, and antenatal prediction. *Obstet Gynecol* 1995; 86:466-75.

Berry SM, Puder KS, Bottoms SF, Uckele JE, Romero R, Cotton DB. Comparison of intrauterine hematologic and biochemical values between twin pairs with and without stuck twin syndrome. *Am J Obstet Gynecol* 1995; 172:1403-10.

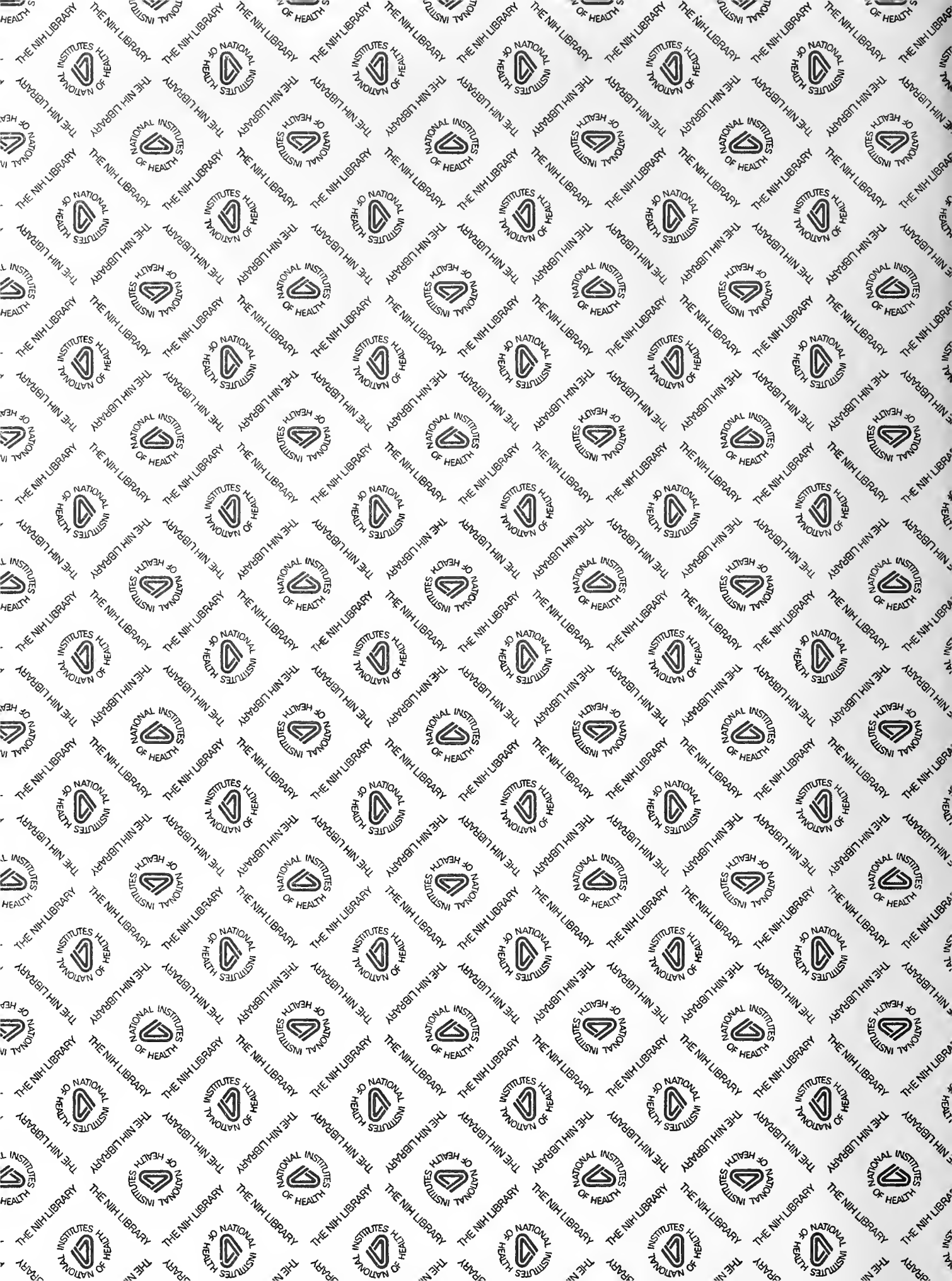
Munoz H, Sherer DM, Romero R, Sanchez J, Hernandez I, Diaz C. Prenatal sonographic findings associated with large cardiac fibroma. *J Ultrasound Med* 1995; 14:479-82.

Puder KS, Sherer DM, Silva M, Ross R, King M, Treadwell MC, Romero R. Prenatal sonographic diagnosis of ductus arteriosus aneurysm. *Ultrasound Obstet Gynecol* 1995; 5(5):342-5.

Quintero, RA, Hume R, Smith C, Johnson, MP, Romero R. Percutaneous fetal cystoscopy and endoscopic fulguration of posterior urethral valves. *Am J Obstet Gynecol* 1995; 172:206-9

Sepulveda W, Pryde PG, Greb AE, Romero R, Evans MI. Prenatal diagnosis of umbilical cord pseudocyst. *Ultrasound Obstet Gynecol* 1994;4:147-50.







Amazing Research.
Amazing Help.

<http://nihlibrary.nih.gov>

10 Center Drive
Bethesda, MD 20892-1150
301-496-1080

